UNEXPLAINED SUBFERTILITY AND ITS INTERVENTIONS

A REAPPRAISAL

NOOR DANHOF

Unexplained subfertility and its interventions; a reappraisal

Nora Alexandra Danhof

Unexplained subfertility and its interventions; a reappraisal

ISBN/EAN: 978-94-6375-777-5

Cover and chapterpage design: Jasper Kroese
Layout and design: Eduard Boxem | www.persoonlijkproefschrift.nl.
Printing: Ridderprint BV | www.ridderprint.nl

Printing of this thesis was supported by Stichting Gynaecologische Endrocrinologie en Kunstmatige Humane Voortplanting, the AMC Medical Research BV, ABN Amro Bank, and efr-software.com

Unexplained subfertility and its interventions; a reappraisal

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K.I.J. Maex

ten overstaan van een door het College voor Promoties ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op vrijdag 6 maart 2020, te 10.00 uur

door

Nora Alexandra Danhof

geboren te Leiden

PROMOTIECOMMISSIE

Promotores:	Prof. dr. F. van der Veen	AMC-UvA
	Prof. dr. S. Repping	AMC-UvA
Copromotores:	Dr. M.H. Mochtar	AMC-UvA
	Dr. M. van Wely	AMC-UvA
Overige leden:	Prof. dr. M.J.C. Eijkemans	Universiteit Utrecht
	Prof. dr. M. Goddijn	AMC-UvA
	Prof. dr. C.B. Lambalk	Vrije Universiteit Amsterdam
	Prof. dr. J.S.E. Laven	Erasmus Universiteit Rotterdam
	Prof. dr. V. Mijatovic	Vrije Universiteit Amsterdam

Faculteit der Geneeskunde

CONTENTS

1.

General Introduction

GENERAL INTRODUCTION

Unexplained subfertility, defined as a proven ovulation, normal semen analysis and open fallopian tubes occurs in forty to fifty percent of all couples seeking medical help because of failure to conceive (Evers, 2002). In the Netherlands, it is estimated that about 4800 couples are diagnosed as such annually (CBS; NHG subfertility; Evers, 2002). Since the pathophysiological mechanism of "unexplained subfertility" is unknown, there is no causal treatment for this condition. By lack of anything better and perhaps because they are the tools of the trade, pragmatic treatment options like intrauterine insemination (IUI) with ovarian stimulation and in vitro fertilization (IVF) are standard clinical practice.

Pharmaceutical drugs which stimulate the ovaries are gonadotrophins, clomiphene citrate or Letrozole. Gonadotrophins stimulate the development of immature follicles in the ovary directly (https://www.drugbank.ca/drugs/DB00066 2019). Clomiphene citrate is a selective estrogen receptor modulator, that inhibits estrogen receptors in the hypothalamus, inhibiting negative feedback of estrogen on gonadotropin release, leading to up-regulation of the hypothalamic-pituary-gonadal axis (DrugBank > Clomifene Archived 2011-06-27 at the Wayback Machine. Updated on April 19, 2011). Letrozole is a nonsteroidal, selective aromatase inhibitor and has therefore an anti-estrogenic effect. Ovarian stimulation carries a risk of multiple pregnancies, which is associated with maternal morbidity, neonatal morbidity and neonatal mortality (Calhaz-Jorge *et al*., 2017; Ombelet et al 2006). Nevertheless, IUI is commonly combined with ovarian stimulation, as it is more effective than unstimulated IUI (Veltman-Verhulst et al., 2012).

IVF bypasses several steps in natural conception, such as transport of spermatozoa through the cervical canal and transport of the fertilized oocytes through the fallopian tubes. IVF is offered to couples with unexplained subfertility, based upon the concept that IVF bypasses unknown pathology in the chain of natural conception. In contrast to IUI with ovarian stimulation, IVF does not carry an increased risk of multiple pregnancies, provided IVF is combined with single embryo transfer (SET) (Practice committee of the American Society of Reproductive Medicine. 2012). The disadvantage of IVF is that it is more invasive than IUI.

The selection of couples with unexplained subfertility for IUI with ovarian stimulation or IVF varies around the world. The Dutch Society of Obstetrics and Gynaecology (NVOG) selects couples based on their predicted chances of natural conception. Couples are offered fertility treatment after twelve months of trying to conceive, provided their chances of natural conception are below 30% in the next twelve months according to Hunault prediction model. The Hunault model integrates the variables female age, duration of subfertility, whether subfertility is primary or secondary, referral status and percentage of motile sperm (Hunault et al 2004). First line treatment is a maximum of six cycles of IUI with ovarian stimulation and in case of failure IVF is offered as second

line treatment (NVOG 2010). The American Society for Reproductive Medicine does advise to consider simple treatment before complex treatment and to start with IUI as a first line treatment, before considering IVF. This society does not mention specifically when to start and acknowledges the need for evidence on the effectiveness of IUI (The Practice Committee of the American Society for Reproductive Medicine., 2006). The Royal College of Obstetricians and Gynaecologists states that IUI may be offered before IVF, as it is cheaper and less invasive, but has doubts on the success rates (RCOG 2019, https://www.rcog.org.uk/en/patients/fertility/treatment/iui/). In contrast, the NICE guideline recommends to not routinely offer IUI, either with or without ovarian stimulation in couples with unexplained subfertility, because of a lack of evidence on the effectiveness of IUI, but advises to try and conceive for a total of 24 months and then to consider IVF (NICE 2017).

So, all in all, the recommendations of the learned societies are based upon low levels of evidence as acknowledged by the societies themselves.

BACKGROUND OF THIS THESIS

In 1998, the first attempt to assess the value of IUI without and with ovarian stimulation and IVF in couples with unexplained subfertility was done in a –now classic- retrospective analysis comparing cost-effectiveness of these interventions. IUI appeared not to be efficacious without some form of ovarian stimulation, and IUI with clomiphene citrate appeared to be more cost-effective than IUI with gonadotrophins or IVF. It was, based upon this analysis, that the provisional recommendation was made that IUI with clomiphene citrate should be tried for several cycles as the initial treatment after expectant management (Guzick et al., 1998). Following up on this study, in another landmark study, 258 couples with unexplained subfertility were randomly allocated to a maximum of six cycles of IUI without ovarian stimulation, six cycles IUI with ovarian stimulation with gonadotrophins or six cycles of IVF. IUI treatment was as effective as IVF, but ovarian stimulation in IUI did not yield higher pregnancy rates. Based upon this RCT, IUI without and with ovarian stimulation seemed to be more cost-effective than IVF. In view of the increased multiple pregnancies after IUI with ovarian stimulation, it was suggested to perform IUI without ovarian stimulation (Goverde et al., 2000).

At the time we realized that these major studies overlooked the issue of prognosis of conception, since this might affect the success rates of interventions. In 2000 the Centre of Reproduction Medicine of Academic Medical Centre started a research line on unexplained subfertility based on prognosis by addressing three major issues. First, the external validity and clinical value of the prediction model of Hunault were evaluated and it appeared that the model also performed well in other populations than the one in which it was developed and was thus valuable for clinical practice. Following this research, the model of Hunault was implemented in the Dutch Society of Obstetrics

and Gynaecology (NVOG) guideline (Prediction models in reproductive medicine. Jan Willem van der Steeg 2008).

Second, to establish which couples with unexplained subfertility would benefit from IUI with ovarian stimulation over expectant management, the model of Hunault was applied in couples with unexplained subfertility and a calculated intermediate prognosis according to Hunault, ie 30%-40% chance of natural conception in the next 12 months. Couples with this intermediate prognosis were randomized between IUI with ovarian stimulation and expectant management. Of the 127 couples that were assigned to IUI with ovarian stimulation, 29 (23%) conceived leading to ongoing pregnancy and of the 126 couples that were assigned to expectant management 34 (27%) conceived leading to ongoing pregnancy (RR 0.85, 95% CI 0.63 to 1.1) (Intrauterine insemination: a reappraisal. Pieternel Steures 2008). Inspired by the data obtained in this trial, a three arm parallel group, pragmatic randomised controlled trial was performed in the UK to compare the effectiveness of clomiphene citrate and unstimulated intrauterine insemination with expectant management for the treatment of unexplained infertility among couples at least two years of subfertility. The duration of subfertility was used as a proxy for prognosis. Also in this trial the interventions did not offer superior live birth rates compared with expectant management (live birth rates: expectant management 17%, clomiphene citrate 14%, unstimulated IUI 23%, clomiphene citrate compared with expectant management, OR 0.79, 95% CI 0.45 to 1.38, unstimulated IUI compared with expectant management OR 1.46, 95% CI 0.88 to 2.43) (Bhattacharya et al., 2008).

Third, a prognostic model was developed for the outcome of IUI. Independent predictors for the chance of an ongoing pregnancy after IUI with ovarian stimulation were maternal age, duration of subfertility, presence of cervical or male factor, presence of one-sided tubal pathology, uterine anomalies, endometriosis, use of ovarian stimulation and IUI cycle number. In an external validation this model appeared to be able to make a good distinction between couples with a good pregnancy chance and those with a poor pregnancy chance after IUI with ovarian stimulation (Intrauterine insemination: a reappraisal. Pieternel Steures 2008, Intra-uterine insemination fine tuning a treatment. Inge Custers 2012). We next focussed on the specific issue of IUI cycle number, as this was of predictive value and because a small cohort study (n=685) had shown that the clinical pregnancy rate per cycle was significantly lower in the second three cycles compared to the first three cycles (Aboulghar et al., 2001). We then performed a much larger retrospective multicentre cohort study among 3714 women, in whom the mean ongoing pregnancy rate was 5.6% per cycle. The ongoing pregnancy rate in the seventh, eighth and ninth cycle were 5.1%, 6.7% and 4.6%, respectively (Intra-uterine insemination fine tuning a treatment. Inge Custers 2012).

Based on these data several conclusions were drawn. First, the prediction model of natural conception by Hunault is of clinical value. Second, initial expectant management for six months is justified in couples with an intermediate prognosis of natural conception according to the model of Hunault. Third, couples with a good pregnancy

chance can be distinguished from those with a poor pregnancy chance after IUI with ovarian stimulation and finally, when couples with unexplained subfertility do start with IUI and ovarian stimulation, the recommendation to limit the number of IUI cycles to six could well be changed into an extension up to nine cycles, although – obviously- pregnancy chances per cycle are the highest in the first three cycles.

In the period 2007-2011 the long-term impact of an initial expectant management on pregnancy rates was studied and it appeared that three years after randomization between IUI with ovarian stimulation and expectant management, during which the recruited couples were managed according to local protocols which included IVF, IUI with ovarian stimulation and expectant management, there were no differences in time to ongoing pregnancy. The women in the initial expectant management arm underwent significantly less IVF cycles and IUI cycles compared to the women in the initial IUI with ovarian stimulation arm (Intra-uterine insemination fine tuning a treatment. Inge Custers 2012).

It was concluded that IUI with ovarian stimulation was unlikely to be effective, also on the long term, in women with unexplained subfertility and an intermediate prognosis on natural conception, while IUI obviously generated costs and carried the risk of multiple pregnancies.

In the period 2006-2017, as a consequence of the newly acquired insight that IUI with ovarian stimulation was not effective in women with unexplained subfertility and an intermediate prognosis on natural conception, the research line concentrated on couples with unexplained subfertility and a poor prognosis of natural conception. The core focus of the research line of the Centre of Reproduction Medicine of Academic Medical Centre was therefore not only to investigate the effectiveness of treatment in these women, but also to investigate safety, ie the risks of multiple pregnancies. First, a systematic review and meta-analysis of twelve cohort studies and two randomized controlled trials suggested that ovarian stimulation in IUI should not aim for more than two dominant follicles per cycle (Outcome measures in reproductive medicine trials. Minouche van Rumste 2013).

Second, in a secondary analysis of individual patient data of randomized controlled trials on treatment strategies for couples with unexplained subfertility, the possibility of using prognosis of natural conception to select the best treatment strategy for these couples was explored. Included treatment strategies were timed intercourse, intracervical insemination and IUI - all without or with ovarian stimulation- and IVF. Data from 8 primary studies were collected, which included 2 550 couples. The results showed that the probability of an ongoing pregnancy after treatment tended to be higher in couples with a better prognosis of natural conception, but a statistically significant differential effect of prognosis of natural conception on treatment efficacy was not found (Tailored expectant management in reproductive medicine. Noortje van den Boogaard 2013).

Third, couples with unexplained subfertility and a poor prognosis of natural conception according to Hunault were recruited for two multicentre randomized trials to establish live birth rates and multiple pregnancy rates comparing IUI with ovarian stimulation to IVF-SET (SETI study and INeS study). IVF-SET was not superior to IUI with ovarian stimulation in terms of effectiveness, safety and costs (Intra-uterine insemination fine tuning a treatment. Inge Custers 2012; The enigma of unexplained subfertility. Alexandra Bensdorp; Unexplained subfertility Illuminating the path to treatment 2016. Raïssa Tjon-Kon-Fat 2017).

Fourth, a secondary analysis of the SETI study showed that a low total motile sperm count tended to lead to higher pregnancy rates after IVF than after IUI with ovarian stimulation. This finding was not confirmed in a second secondary analysis of the – much larger- INeS study (Unexplained subfertility Illuminating the path to treatment. Raïssa Tjon-Kon-Fat 2017).

Based on these studies it was concluded that IUI with ovarian stimulation indeed should be the first line treatment for couples with unexplained subfertility and a poor prognosis of natural conception and that this intervention should probably aim for not more than two follicles per cycle.

In the period 2012-2019 the focus was placed on prognosis and treatment of couples with unexplained subfertility with either IUI or IVF in the context of time. The major shortcoming of the model of Hunault was that it can only be used at one moment in time, explicitly direct after the fertility work up. First, a new dynamic prediction model which accounts for the principle of selection over time when calculating prognoses was developed and externally validated (Medically assisted reproduction in the context of time. Irma Scholten 2015, Prognosis-based management of couples with unexplained subfertility. Rik van Eekelen 2019).

Second, in a prospective cohort study, couples with unexplained subfertility after expectant management were followed in which a subset of couples started IUI with ovarian stimulation at various points in time. There were 800 couples who at least had one cycle of IUI with ovarian stimulation within 1.5 years post fertility workup of whom 142 couples conceived (ongoing pregnancy rate: 0.50 per couple per year, median follow up 4 months). Out of 1096 untreated couples, 386 conceived naturally (ongoing pregnancy rate: 0.31 per couple per year, median follow up 7 months) (Prognosis-based management of couples with unexplained subfertility. Rik van Eekelen 2019). In the intervening time, the TUI trial was published, in which couples with unexplained subfertility and a poor prognosis of natural conception were randomized between IUI with ovarian stimulation and expectant management. This trial showed that IUI with ovarian stimulation significantly increased cumulative live birth rates compared to expectant management (OR with 95% CIs: 3.4, 1.7–6.8) (Farquhar et al., 2017).

Third, in an observational study pregnancy outcomes were compared between couples undergoing IVF (n=40921) obtained from British registry data and couples undergoing

expectant management comprised from a prospective nation-wide Dutch cohort (n=4875) and a retrospective regional cohort from Scotland (n=975). The adjusted chance of conception was 47.9% (95% CI 45.0 to 50.0) after IVF and 26.1% (95% CI 24.2 to 28.0) after expectant management. The average absolute difference in the adjusted one year chance of conception was 21.8% (95% CI 18.3 to 25.3) (Prognosis-based management of couples with unexplained subfertility. Rik van Eekelen 2019).

There were two conclusions drawn from the research in this period. First, the new dynamic prediction model allowed to better investigate the value of both IUI with ovarian stimulation and IVF over expectant management in couples with unexplained subfertility and a poor prognosis of natural conception, as it takes into account the profound effect of time. Second, IUI with ovarian stimulation seemed to be more effective than expectant management in couples with unexplained subfertility and a poor prognosis of natural conception.

In summary, research on treatment modalities in unexplained subfertility over the last two decades has shown that expectant management for at least six months is justified in couples with an intermediate prognosis of natural conception. The prediction model on natural conception by Hunault has been improved by the development of a dynamic prediction model of that is able to give repeated predictions. In couples with a poor prognosis of natural conception, IVF-SET is not more cost-effective than IUI with ovarian stimulation. Based on cohort studies and one randomized controlled trial, both IUI with ovarian stimulation and IVF-SET improve pregnancy rates over expectant management in couples who have a poor prognosis of natural conception and the benefit of these interventions increases when the prognosis of natural conception declines.

Most data generated by the long standing research line at the Centre for Reproductive Medicine of the Academic Medical Centre of Amsterdam on treatment modalities for unexplained subfertility have been integrated in the multi-year project of the World Health Organization to review the evidence for the establishment of normative guidance for the implementation of IUI as a treatment to address fertility problems (Cohlen et al., 2018).

This thesis addresses two remaining knowledge gaps.

First, the basic position of most clinical guidelines for the management of unexplained subfertility is to recommend starting with the least invasive intervention before moving on to interventions that are more invasive, but this is not based on well-founded evidence (ASRM 2006; NICE 2017; NVOG 2010; RCOG 2019). To date, several randomized controlled trials have investigated the effectiveness and safety of IUI with ovarian stimulation, IVF-SET and expectant management, but only in separate head to head comparisons. These studies have never been integrated into one analysis to provide well-founded recommendations.

Second, safety of IUI with ovarian stimulation is still a clinical challenge. In 2007, a randomized controlled trial was performed to compare ovarian stimulation with gonadotrophins to ovarian stimulation with clomiphene citrate in IUI for unexplained subfertility whereby the number of dominant follicles per cycle was limited through withholding insemination when more than three dominant follicles developed. The result was a low multiple pregnancy rate (2,2%) without compromising effectiveness (28%), but uncertainty remained as this study was underpowered with 138 women included (Dankert et al., 2007). Nonetheless, IUI with ovarian stimulation was still practiced without considering the number of dominant follicles. In a large RCT among 900 couples with unexplained subfertility, IUI with gonadotrophins was compared to IUI with clomiphene citrate and IUI with Letrozole. IUI with gonadotrophins showed a statistically significant increase in live birth rates (32%) compared to clomiphene citrate (23%) and Letrozole (19%), but at the cost of 25 twins and six triplets among 301 women (10%) undergoing ovarian stimulation with gonadotrophins, while there were 8 twins among 300 women (3%) undergoing ovarian stimulation with CC and there were 6 twins among 299 women (2%) (Diamond et al., 2015). These high multiple pregnancy rates are no longer acceptable in modern infertility treatment.

Although a strategy with adherence to strict cancellation criteria seems the regimen of choice, it is unclear which pharmaceutical agent is the most effective and safe, whether there are specific women at baseline that benefit from one or the other pharmaceutical agent and whether the impact of the pharmaceutical agent on endometrial thickness is associated with ongoing pregnancy. It is also unclear how the costs of pharmaceutical agents are related to the effectiveness.

OUTLINE OF THE THESIS

Chapter 2 presents the results of a network meta-analysis on the effectiveness of treatment modalities for unexplained subfertility including expectant management as well as the active treatment modalities: ovarian stimulation (OS), intrauterine insemination (IUI), OS-IUI and in-vitro fertilisation (IVF)/Intracytoplasmic sperm injection (ICSI).

Chapter 3 reports the results of a randomised controlled trial comparing gonadotrophins to clomiphene citrate in women with unexplained subfertility undergoing IUI of which the quintessence is the adherence to strict cancellation criteria.

Chapter 4 presents the results of a network meta-analysis on the effectiveness and safety of 4 strategies of IUI for unexplained subfertility, namely IUI with ovarian stimulation with gonadotrophins, clomiphene citrate or Letrozole or IUI in a natural cycle

Chapter 5 explores whether it is possible to identify women at baseline that would have better chances of an ongoing pregnancy with one or the other pharmaceutical agent.

Chapter 6 evaluates the impact of endometrial thickness after ovarian stimulation with gonadotrophins or clomiphene citrate in the context of IUI and the association between endometrial thickness and the ongoing pregnancy rate.

Chapter 7 presents a cost-effectiveness analysis alongside the randomised controlled trial discussed in chapter 3.

Chapter 8 presents a general discussion of this thesis and implications for clinical practice and further research.

Chapter 9 presents the summary of all chapters.

Chapter 10 presents the Dutch translation of the summary presented in chapter 9.

REFERENCES

Aboulghar M, Mansour R, Serour G, Abdrazek A, Amin Y, Rhodes C. Controlled ovarian hyperstimulation and intrauterine insemination for treatment of unexplained infertility should be limited to a maximum of three trials. Fertil Steril 2001;75:88–91.

Bhattacharya S, Harrild K, Mollison J, Wordsworth S, Tay C, Harrold A, McQueen D, Lyall H, Johnston L, Burrage J et al. . Clomifene citrate or unstimulated intrauterine insemination compared with expectant management for unexplained infertility: pragmatic randomised controlled trial. Br Med J 2008;337:a716.

Bensdorp, A.J. The enigma of unexplained subfertility. Thesis 2016

Calhaz-Jorge C, De Geyter C, Kupka MS, de Mouzon J, Erb K, Mocanu E, et al. Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod. 2017;32:1957-73.

CBS Statistics Netherlands, https://www.cbs.nl/

Custers, IM. Intra-uterine insemination fine tuning a treatment. Thesis 2012

Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H and van Dop PA. A randomized clinical trial of clomifene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. Hum Reprod. 2007;3:792-7

Cohlen B, Bijkerk A, Van der Poel S, Ombelet W. IUI: review and systematic assessment of the evidence that supports global recommendations. Human Reproduction Update, Volume 24, Issue 3, May-June 2018, Pages 300–319

Diamond MP, Richar SL, Coutifaris C, Alvero R, Robinson R, Casson P, Christman GM, Ager J, Huang H, Hansen KR et al. Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility. N Engl J Med. 2015;13:1230-40

Evers JL. Female subfertility. *Lancet* 2002;360:151-159.

Farquhar CM, Liu E, Armstrong S, Arroll N, Lensen S, Brown J. Intrauterine insemination with ovarian stimulation versus expectant management for unexplained infertility (TUI): a pragmatic, open-label, randomised, controlled, two-centre trial. Lancet. 2018 Feb 3;391(10119):441-450.

Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. Hum Reprod. 2004 Sep;19:2019-26.

Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. Lancet (London, England) 2000;355:13–18.

Guzick DS, Sullivan MW, Adamson GD, Cedars MI, Falk RJ, Peterson EP, Steinkampf MP. Efficacy of treatment for unexplained infertility. Fertil Steril 1998;70:207–213.

NICE guideline 2017, https://www.nice.org.uk/guidance/cg156/chapter/Recommendation

NVOG guideline: onverklaarde subfertiliteit, 2010, https://www.nvog.nl/wp-content/uploads/2017/12/Onverklaarde-subfertiliteit-1.0-15-09-2010.pdf

Ombelet W, Martens G, De Sutter P, Gerris J, Bosmans E, Ruyssinck G, et al. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod. 2006;21(4):1025-32.

Practice Committee of the American Society for Reproductive Medicine. Effectiveness and treatment for unexplained infertility. Fertil steril. 2006;86(5):111-114

Practice committee of the American Society of Reproductive Medicine. Multiple gestation associated with infertility therapy: an American Society for Reproductive Medicine Practice Committee opinion. Fertil Steril. 2012;97(4):825-34

RCOG 2019, https://www.rcog.org.uk/en/patients/fertility/treatment/iui/

Scholten, I. Medically assisted reproduction in the context of time. Thesis 2015

Steures P. Intrauterine insemination: a reappraisal. Thesis 2008

Tjon-Kon-Fat, R.I. Illuminating the path to treatment. Thesis 2017

Van Asselt KM, Hinloopen RJ, Silvius AM, Van der Linden PJQ, Van Oppen CCAN, Van Balen JAM. NHG-Standaard Subfertiliteit (Tweede herziening). Huisarts Wet 2010:53:203-14.

van Eekelen, R. Prognosis-based management of couples with unexplained subfertility. Thesis 2019.

van den Boogaard, N.M. Tailored expectant management in reproductive medicine. Thesis 2013

van der Steeg, J.W. Prediction models in reproductive medicine. Thesis 2008

van Rumste, M. Outcome measures in reproductive medicine trials. Thesis 2013

Veltman-Verhulst SM, Cohlen BJ, Hughes E, Heineman MJ. Intra-uterine insemination for unexplained subfertility. Cochrane Database Syst Rev 2012;9:CD001838.

2.

Interventions for unexplained infertility: a systematic review and network meta-analysis

authors:
R. Wang, **N.A. Danhof,** R.I. Tjon-Kon-Fat, M.J.C. Eijkemans, P.M.M. Bossuyt, M.H. Mochtar, F. van der Veen, S. Bhattacharya, B.W.J Mol, M. van Wely

Cochrane Database of Systematic Reviews 2019;5:9:CD012692

ABSTRACT

Background: Clinical management for unexplained infertility includes expectant management as well as active treatments, including ovarian stimulation (OS), intrauterine insemination (IUI), OS-IUI, and in vitro fertilisation (IVF) with or without intracytoplasmic sperm injection (ICSI).

Existing systematic reviews have conducted head-to-head comparisons of these interventions using pairwise meta-analyses. As this approach allows only the comparison of two interventions at a time and is contingent on the availability of appropriate primary evaluative studies, it is difficult to identify the best intervention in terms of effectiveness and safety. Network meta-analysis compares multiple treatments simultaneously by using both direct and indirect evidence and provides a hierarchy of these treatments, which can potentially better inform clinical decision-making.

Objectives: To evaluate the effectiveness and safety of different approaches to clinical management (expectant management, OS, IUI, OS-IUI, and IVF/ICSI) in couples with unexplained infertility.

Search methods: We performed a systematic review and network meta-analysis of relevant randomised controlled trials (RCTs). We searched electronic databases including the Cochrane Gynaecology and Fertility Group Specialised Register of Controlled Trials, the Cochrane Central Register of Studies Online, MEDLINE, Embase, PsycINFO and CINAHL, up to 6 September 2018, as well as reference lists, to identify eligible studies. We also searched trial registers for ongoing trials.

Selection criteria: We included RCTs comparing at least two of the following clinical management options in couples with unexplained infertility: expectant management, OS, IUI, OS-IUI, and IVF (or combined with ICSI).

Data collection and analysis: Two review authors independently screened titles and abstracts identified by the search strategy. We obtained the full texts of potentially eligible studies to assess eligibility and extracted data using standardised forms. The primary effectiveness outcome was a composite of cumulative live birth or ongoing pregnancy, and the primary safety outcome was multiple pregnancy. We performed a network meta-analysis within a random-effects multi-variate meta-analysis model. We presented treatment effects by using odds ratios (ORs) and 95% confidence intervals (CIs). For the network meta-analysis, we used Confidence in Network Meta-analysis (CINeMA) to evaluate the overall certainty of evidence.

Main results: We included 27 RCTs (4349 couples) in this systematic review and 24 RCTs (3983 couples) in a subsequent network meta-analysis. Overall, the certainty of evidence was low to moderate: the main limitations were imprecision and/or heterogeneity.

Ten RCTs including 2725 couples reported on live birth. Evidence of differences between OS, IUI, OS-IUI, or IVF/ICSI versus expectant management was insufficient (OR 1.01,

95% CI 0.51 to 1.98; low-certainty evidence; OR 1.21, 95% CI 0.61 to 2.43; low-certainty evidence; OR 1.61, 95% CI 0.88 to 2.94; low-certainty evidence; OR 1.88, 95 CI 0.81 to 4.38; low-certainty evidence). This suggests that if the chance of live birth following expectant management is assumed to be 17%, the chance following OS, IUI, OS-IUI, and IVF would be 9% to 28%, 11% to 33%, 15% to 37%, and 14% to 47%, respectively. When only including couples with poor prognosis of natural conception (3 trials, 725 couples) we found OS-IUI and IVF/ICSI increased live birth rate compared to expectant management (OR 4.48, 95% CI 2.00 to 10.1; moderate-certainty evidence; OR 4.99, 95 CI 2.07 to 12.04; moderate-certainty evidence), while there was insufficient evidence of a difference between IVF/ICSI and OS-IUI (OR 1.11, 95% CI 0.78 to 1.60; low-certainty evidence).

Eleven RCTs including 2564 couples reported on multiple pregnancy. Compared to expectant management/IUI, OS (OR 3.07, 95% CI 1.00 to 9.41; low-certainty evidence) and OS-IUI (OR 3.34 95% CI 1.09 to 10.29; moderate-certainty evidence) increased the odds of multiple pregnancy, and there was insufficient evidence of a difference between IVF/ICSI and expectant management/IUI (OR 2.66, 95% CI 0.68 to 10.43; low-certainty evidence). These findings suggest that if the chance of multiple pregnancy following expectant management or IUI is assumed to be 0.6%, the chance following OS, OS-IUI, and IVF/ICSI would be 0.6% to 5.0%, 0.6% to 5.4%, and 0.4% to 5.5%, respectively.

Trial results show insufficient evidence of a difference between IVF/ICSI and OS-IUI for moderate/severe ovarian hyperstimulation syndrome (OHSS) (OR 2.50, 95% CI 0.92 to 6.76; 5 studies; 985 women; moderate-certainty evidence). This suggests that if the chance of moderate/severe OHSS following OS-IUI is assumed to be 1.1%, the chance following IVF/ICSI would be between 1.0% and 7.2%.

Authors' conclusions: There is insufficient evidence of differences in live birth between expectant management and the other four interventions (OS, IUI, OS-IUI, and IVF/ICSI). Compared to expectant management/IUI, OS may increase the odds of multiple pregnancy, and OS-IUI probably increases the odds of multiple pregnancy. Evidence on differences between IVF/ICSI and expectant management for multiple pregnancy is insufficient, as is evidence of a difference for moderate or severe OHSS between IVF/ICSI and OS-IUI.

PLAIN LANGUAGE SUMMARY

Interventions for unexplained infertility: a systematic review and meta-analysis

Review question

Researchers in Cochrane reviewed the evidence on the effectiveness and safety of ovarian stimulation (OS), intrauterine insemination (IUI), OS-IUI, and in vitro fertilisation (IVF) with or without intracytoplasmic sperm injection (ICSI) versus expectant management in couples with unexplained infertility.

Background

Treatment options for unexplained infertility include expectant management as well as active treatments such as ovarian stimulation (OS), intrauterine insemination (IUI), OS-IUI, and in vitro fertilisation (IVF) with or without intracytoplasmic sperm injection (ICSI). Network meta-analysis synthesises evidence of direct and indirect comparisons of interventions and enables researchers to simultaneously assess the effectiveness of more than two interventions for the same condition, so that clinicians can use the evidence to offer the best treatment. Therefore, we compared all these different treatment options by using network meta-analysis, to better inform clinical decision-making.

Study characteristics

We found 27 randomised controlled trials comparing these treatments with each other in a total of 4349 couples with unexplained infertility. The evidence is current to September 2018.

Key results

Evidence of differences in live birth between expectant management and the other four treatments (OS, IUI, OS-IUI, and IVF/ICSI) was insufficient. If the chance of live birth following expectant management is assumed to be 17%, the chance following OS, IUI, OS-IUI, and IVF would be 9% to 28%, 11% to 33%, 15% to 37%, and 14% to 47%, respectively. Compared to expectant management/IUI, OS may increase the chances of multiple pregnancy, and OS-IUI probably increases the chances of multiple pregnancy. Evidence showing differences between IVF/ICSI and expectant management for multiple pregnancy was insufficient. If the chance of multiple pregnancy following expectant management/IUI is assumed to be 1%, the chance following OS, OS-IUI, and IVF/ICSI would be 1% to 5%, 1% to 5%, and 0% to 6%, respectively.

Certainty of the evidence

The certainty of evidence overall was low to moderate. The main limitations were imprecision (not enough couples have been studied) and heterogeneity (couples in existing studies had different clinical characteristics).

BACKGROUND

Description of the condition

Up to one in eight couples who try to achieve pregnancy fail to do so after 12 months of unprotected intercourse (Boivin, et al., 2007, Datta, et al., 2016, Gnoth, 2003). Routine fertility investigations comprising semen analysis, assessment of ovulation, and a tubal patency test fail to reveal any abnormality in 25% of couples who are said to have unexplained infertility (Brandes, et al., 2010, Hull, et al., 1985). In the absence of an obvious barrier to conception, many of these couples possess a good chance of achieving pregnancy without treatment (Brandes, et al., 2011).

Description of the intervention

Clinical guidelines for the management of unexplained infertility recommend starting with the least invasive intervention before moving on to those that are more invasive (American Society for Reproductive, 2006, Dutch Society of Obstetrics and Gynaecology, 2010, National Institute for Health and Care Excellence, 2013). In clinical practice, this has led to a wide range of clinical management approaches, ranging from expectant management (i.e. sexual intercourse) to timed intercourse, ovarian stimulation (i.e. gonadotropins, aromatase inhibitors, or anti-oestrogens), intrauterine insemination (IUI) with or without ovarian stimulation, in vitro fertilisation (IVF), and intracytoplasmic sperm injection (ICSI).

Expectant management or timed intercourse

Couples have a good chance of achieving pregnancy without treatment. A cumulative ongoing pregnancy rate of 27% has been reported after 12 months of unprotected intercourse following completion of the fertility investigations (Hunault, et al., 2005, van Eekelen, et al., 2017)

Ovarian stimulation (OS)

Anti-oestrogens (e.g. clomiphene), gonadotropins (e.g. urinary or recombinant follicle-stimulating hormone), and aromatase inhibitors (e.g. letrozole) are the most commonly used medications for OS. OS is used to stimulate follicular growth to increase the number of mature oocytes available for fertilisation, assuming that this would increase the chance of a live birth.

IUI (with or without OS)

IUI is another treatment option for unexplained infertility. It involves placement of prepared sperm into the uterine cavity timed around ovulation (Kandavel and Cheong, 2018). IUI can be done in a natural cycle or in combination with OS. Live birth rates of approximately 6% to 10% per cycle have been reported for infertile couples with unexplained infertility undergoing IUI with or without ovarian stimulation (Huang, et al., 2018).

IVF and ICSI

Conventional IVF refers to the co-incubation of oocytes with sperm in vitro with the goal of achieving extracorporeal fertilisation (Zegers-Hochschild, et al., 2017); this was first used as a treatment option for tubal infertility (Steptoe and Edwards, 1978). ICSI is a procedure in which a single spermatozoon is injected into the oocyte cytoplasm (Zegers-Hochschild, et al., 2017); this was first used in couples with severe male factor infertility (Palermo, et al., 1992). In the last three decades, the indication for IVF and ICSI has expanded to embrace a wider range of couples with infertility, including those with unexplained infertility (Kamphuis, et al., 2014).

How the intervention might work

In couples with unexplained infertility, a biological cause for their involuntary childlessness has not been detected, and therefore the rationale for each possible treatment is based upon assumptions.

The concept behind timed intercourse is to aid couples in having intercourse at the best time for fertilisation through the use of cycle monitoring. Ovarian stimulation is used to stimulate follicular growth to increase the number of mature oocytes available for fertilisation. IUI brings the spermatozoa closer to the oocyte for fertilisation at the appropriate time. The combination of OS and IUI combines these effects. IVF bypasses the process of transport of spermatozoa. ICSI assists fertilisation in overcoming any subtle abnormalities of sperm-oocyte interaction.

Why it is important to do this review

Various reviews have examined interventions for couples with unexplained infertility (Athaullah, et al., 2002, Gunn and Bates, 2016, Hughes, et al., 2010, Pandian, et al., 2015, Veltman-Verhulst, et al., 2016). These reviews have included head-to-head comparisons of two interventions. Given that no large randomised controlled trials (RCTs) have compared all these available treatments, it is still uncertain which one is the most effective and safe option. Network meta-analysis could synthesise and interpret the wider picture of existing evidence by incorporating both direct and indirect evidence of different interventions. This approach can also identify gaps in research that need to be addressed in the future.

Objectives

To evaluate the effectiveness and safety of different approaches of clinical management (expectant management, OS, IUI, OS-IUI, and IVF/ICSI) in couples with unexplained infertility.

METHODS

Criteria for considering studies for this review

Types of studies

All randomised controlled trials (RCTs) comparing the effectiveness and/or safety of one of the interventions versus the other intervention. We excluded quasi-randomised and non-randomised studies. Cross-over trials were included, but only data from the first phase were used.

Types of participants

Couples who had been trying to conceive for at least one year, women having at least one patent fallopian tube and an ovulatory cycle, and men having a pre-wash total motile sperm count > 3 * 10ˆ6 were eligible. Among women with a diagnosis of endometriosis, only those with mild endometriosis (American Fertility Society (AFS) criteria I) were included.

Types of interventions

We considered all trials that compared at least two of the following clinical management options.

- Expectant management, including timed intercourse.
- OS using gonadotropins, aromatase inhibitors, anti-oestrogens, or their combination.
- IUI without ovarian stimulation.
- OS-IUI.
- IVF with a single embryo transfer, with a double embryo transfer, or combined with ICSI.

Expectant management and timed intercourse were combined in the same group if no invasive techniques were used. Studies comparing different OS protocols were excluded and those comparing OS with different protocols were pooled as one OS group. The five proposed interventions were jointly randomisable (i.e. a couple with unexplained infertility is theoretically able to be randomised to any of the five interventions). ICSI was not considered as a separate intervention, as it is indicated for couples with severe male factor infertility or with fertilisation failure in previous IVF cycles. Therefore, ICSI was not jointly randomisable with the other interventions and including ICSI will violate the transitivity assumption in this network meta-analysis. Moreover, trials including IVF as an intervention often also applied ICSI for couples with unexpected low sperm count on the day of oocyte retrieval, or with previous IVF failure in a multi-cycle intervention; therefore IVF with and without ICSI was considered as the same intervention. Studies with an embryo transfer policy allowing transfer of more than two embryos in an unselected population were included in the systematic review but were excluded from the network meta-analysis to make the transitivity assumption valid. Natural cycle IVF

and modified natural cycle IVF were not included, as they are not comparable to other IVF protocols.

Types of outcome measures
Primary outcomes

- The primary effectiveness outcome was a composite of cumulative live birth or ongoing pregnancy per woman randomised. Live birth was defined as the birth of a living child after 24 weeks of gestation. Ongoing pregnancy was defined as at least one registered embryonic heartbeat on ultrasound at 12 weeks' gestation and was used in the analysis only when live birth was not reported. Cumulative refers to multiple attempts to achieve a live birth (i.e. multiple cycles of treatments). In IVF, cumulative refers to fresh embryo transfer followed by frozen embryo transfer cycles when applicable
- The primary safety outcome was multiple pregnancy per woman randomised (defined as at least two registered embryonic heartbeats on ultrasound)

Secondary outcomes

- Clinical pregnancy per woman randomised (defined as at least one registered embryonic heartbeat on ultrasound)
- Moderate/severe ovarian hyperstimulation syndrome (OHSS) per woman randomised (defined as moderate abdominal pain, nausea ± vomiting, the presence of ascites on ultrasound or clinical ascites, and ovarian size of at least 8 cm) (Mathur, et al., 2005)

Search methods for identification of studies

We searched for all published and unpublished RCTs, without language or date restrictions, in consultation with the Cochrane Gynaecology and Fertility Group (CGF) Information Specialist.

Electronic searches
We searched the following electronic databases for relevant trials.

- The Cochrane Gynaecology and Fertility Group (CGF) Specialised Register of Controlled Trials, searched 6 September 2018 (Procite platform) (Appendix 1).
- The Cochrane Central Register of Studies Online, searched 6 Sptember 2018 (CRSO Web platform) (Appendix 2).
- MEDLINE, searched from 1946 to 6 September 2018 (Ovid platform) (Appendix 3).
- Embase, searched from 1980 to 6 September 2018 (Ovid platform) (Appendix 4).
- PsycINFO, searched from 1806 to 6 September 2018 (Ovid platform) (Appendix 5).
- Cumulative Index to Nursing and Allied Health Literature (CINAHL), searched from 1961 to 6 September 2018 (Ebsco platform) (Appendix 6).

The MEDLINE search was combined with the Cochrane highly sensitive search strategy for identifying randomised trials, which appeared in the *Cochrane Handbook for Systematic Reviews of Interventions*(Version 5.1.0, Chapter 6, 6.4.11). Embase, PsycINFO, and CINAHL searches were combined with trial filters developed by the Scottish Intercollegiate Guidelines Network (www.sign.ac.uk/methodology/filters.html#random).

Other electronic sources of trials will include the following.

- Trial registers for ongoing and registered trials.
  - o www.clinicaltrials.gov (a service of the US National Institutes of Health).
  - o www.who.int/trialsearch/Default.aspx (the World Health Organization International Trials Registry Platform search portal).
- Virtual Health Library Regional Portal (VHL) (bvsalud.org/portal/?lang=en), which includes Latin American Caribbean Health Sciences Literature (LILACS).
- PubMed and Google Scholar (for recent trials not yet indexed in the major databases).

Searching other resources

We handsearched the reference lists of relevant trials and systematic reviews retrieved by the search and contacted experts in the field to obtain additional data. We also handsearched relevant journals and conference abstracts that were not covered in the CGFG Register, in liaison with the Information Specialist.

Data collection and analysis

Selection of studies

At least two review authors (from RW, RIT, NAD) independently assessed trial eligibility, according to the Criteria for considering studies for this review. We resolved disagreements through discussion with another review author (MvW). We drew a PRISMA flow diagram to show the results of the search and the numbers of included and excluded trials. Reasons for excluding from the (network) meta-analysis any potentially eligible studies identified by the search were documented.

Data extraction and management

For all included trials, two review authors (RW, NAD) independently extracted data using a data abstraction form and summarised trial characteristics in tables. From each included study, two review authors (RW, NAD) extracted baseline characteristics of couples, study settings, methods, types of interventions (used dose, type of preparation, regimen, co-interventions), and outcomes. We intended to contact the study investigators for further data on methods and results, if required.

Assessment of risk of bias in included studies

Two review authors (RW, NAD) independently assessed risk of bias for each eligible study by using the Cochrane 'Risk of bias' assessment tool (Higgins and Green, 2011),

which included six domains: selection (random sequence generation and allocation concealment); performance (blinding of participants and personnel); detection (blinding of outcome assessors); attrition (incomplete outcome data); reporting (selective reporting); and other bias. Disagreements were resolved by discussion with a third review author (MvW). We described all judgements fully and presented our conclusions in the 'Risk of bias' table, which we incorporated into the interpretations of review findings by performing sensitivity analyses.

Measures of treatment effect

As all outcomes involved dichotomous data, we used the numbers of events in control and intervention groups of each study to calculate Mantel-Haenszel odds ratios (ORs). We presented 95% confidence intervals (CIs) for all outcomes. Furthermore, we calculated the probability that an intervention was ranked first, second, and so on. We displayed this ranking graphically in cumulative rankograms for the primary and secondary outcomes using the surface under the cumulative ranking (SUCRA), where SUCRA values can range from zero (i.e. the intervention is certain to be the worst) to one (i.e. the intervention is certain to be the best) (Salanti, et al., 2011).

Unit of analysis issues

The primary unit of analysis was cumulative rates for each outcome per woman randomised. Multiple births were counted as one live birth event. Only first-phase data from cross-over trials were included. Trials comparing the same number of cycles/ months of expectant management, OS, IUI, and OS-IUI were included. As one cycle of IVF takes longer than the other treatments, studies comparing the same cycles of IVF and other treatments were not included in the network meta-analysis but were included in the systematic review. Trials comparing IVF and other treatments within the same period of time were included in the network meta-analysis.

Dealing with missing data

We analysed the data on an intention-to-treat basis as far as possible (i.e. including all randomised participants in the analysis, in the groups to which they were randomised). We attempted to obtain missing data from existing Cochrane Reviews or from the original trialists. If data could not be obtained, we assumed the missing values as a non-event outcome and undertook imputation of individual values only for the primary outcome. For other outcomes, we analysed only available data. Any imputation undertaken was subjected to sensitivity analysis.

Assessment of heterogeneity

Clinical and methodological heterogeneity

To identify clinical and methodological heterogeneity, we compared descriptive statistics for trial and study population characteristics across all eligible trials comparing each pair of interventions. Additionally, we considered whether there was sufficient similarity in the studied interventions and the characteristics of couples across all included studies for inclusion in the network meta-analysis (i.e. the assumption of transitivity in network

meta-analyses). We explored the distribution of potential effect modifiers across various interventions (i.e. female age, and duration of infertility). In this study, we expected the transitivity assumption to hold true assuming the following.

- The nature of the common intervention used for indirect comparisons was consistent (e.g. IUI in an RCT comparing IUI with expectant management was the same as IUI in an RCT comparing IUI with IVF/ICSI).
- All pairwise comparisons did not differ with respect to the distribution of effect modifiers (e.g. design and study characteristics of an RCT comparing IUI vs expectant management were similar to those of an RCT comparing IUI vs IVF/ICSI).

Statistical heterogeneity and inconsistency

Within each pairwise comparison, we assessed statistical heterogeneity by using the I^2 statistic. An I^2 value greater than 50% was taken as an indication of substantial heterogeneity (Higgins and Green, 2011).

In the network meta-analysis, we assessed inconsistency in the network through two approaches: the design-by-treatment method for global approach (Higgins, et al., 2012), and the side-splitting method for local approach (Dias, et al., 2010). The design-by-treatment interaction model allowed for global statistical testing for the presence of inconsistency in the whole network (Higgins, et al., 2012). The local approach identified disagreements between direct and indirect comparisons within each comparison within closed loops in the network (Dias, et al., 2010).

Assessment of reporting biases

In view of the difficulty of detecting and correcting for publication bias and other reporting biases, we aimed to minimise their potential impact by ensuring a comprehensive search for eligible studies and by being alert for duplication of data. If we included ten or more studies in an analysis, we used a comparison-adjusted funnel plot to explore the possibility of small study effects (a tendency for estimates of the intervention effect to be more beneficial in smaller studies) (Chaimani, et al., 2013).

Data synthesis

We compared interventions using odds ratios (ORs) with their respective 95% confidence intervals (CIs). If more than two studies compared the same treatments, a random-effects summary OR was calculated in a pairwise meta-analysis.

We conducted a network meta-analysis based on all investigated comparisons between treatments, in which the indirect analysis was performed by utilising all pathways within the network. An indirect estimate of A versus B can be calculated by comparing direct comparisons of A versus C with comparisons of B versus C. In this way, the OR for comparing A and B can be calculated using the following principle: $\ln(OR_{AvsB}) = \ln(OR_{AvsC}) - \ln(OR_{BvsC})$. We performed a frequentist network meta-analysis within a random-effects multi-variate meta-analysis model (White, 2015).

We assumed a common estimate for the heterogeneity variance across the different comparisons. We used Review Manager (version 5.3, The Cochrane Collaboration) for pairwise meta-analyses and Stata software (version 15.1, Statacorp) for network meta-analyses (Chaimani and Salanti, 2015, White, 2015).

Subgroup analysis and investigation of heterogeneity

If data were available from at least two studies, we conducted subgroup analyses for the primary outcomes only to determine the separate evidence within the following subgroups.

- Women aged ≦ 38 years versus women aged > 38 years.
- Short duration of infertility (≦ 2 years) versus long duration of infertility (> 2 years).
- IVF/ICSI with single embryo transfer policy and IVF/ICSI with non-single embryo transfer policy.

Sensitivity analysis

We conducted sensitivity analyses for live birth/ongoing pregnancy to determine whether the conclusions were robust to arbitrary decisions made regarding eligibility and analysis. These analyses included consideration of whether the review conclusions would have differed if:

- eligibility had been restricted to studies with no domains at high risk of bias;
- alternative imputation strategies had been implemented;
- eligibility had varied by publication type (abstract vs full text); or
- only studies with the outcome live birth had been included.

Overall certainty of the body of evidence: 'Summary of findings' table

We presented overall certainty of the body of evidence for the main review outcomes for each comparison in 'Summary of findings' tables. We evaluated the overall certainty of the evidence based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach in line with a framework developed by Salanti and colleagues in an online tool - Confidence in Network Meta-analysis (CINeMA) (CINeMA, 2017, Salanti, et al., 2014). Domains included study limitations (risk of bias), inconsistency, imprecision, indirectness, and publication bias. For study limitations, we incorporated the contribution of each direct estimate into the overall network estimate when making judgements of study limitations. As blinding was not possible due to the nature of the interventions, we did not downgrade overall certainty if performance bias was the only issue in study limitations. For inconsistency, we evaluated both between-study heterogeneity and disagreements between direct and indirect evidence (i.e. incoherence). We evaluated heterogeneity by considering the agreement of conclusions based on confidence and prediction intervals in relation to the clinically important effect size, in which the major consideration was whether heterogeneity impacts clinical

decisions. If heterogeneity (presented in a prediction interval) impacted decision-making based on a confidence interval, we downgraded the certainty of evidence. We evaluated incoherence by assessing local and global inconsistency. For comparisons with local inconsistency, we downgraded the level of certainty in relevant comparisons. Judgements about evidence certainty (high, moderate, low, or very low) were justified, documented, and incorporated into the reporting of results for each outcome.

RESULTS

Description of studies

Results of the search

The initial electronic database search yielded 2095 articles, with nine additional articles identified through handsearches or searches of trial registers. After removing duplicates, we screened 1171 studies. Screening of titles and abstracts led to the exclusion of 1111 irrelevant studies; 60 full-text articles were further assessed for eligibility. Another 23 studies were further excluded, including five ongoing studies (NCT01992731, 2013, NCT02001870, 2013, NCT02461173, 2015, NCT03455426, 2018, NTR5599, 2016). Finally, 27 studies fulfilled the inclusion criteria as shown in Figure 1 (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Arici, et al., 1994, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Crosignani, et al., 1991, Custers, et al., 2011, Deaton, et al., 1990, Elzeiny, et al., 2014, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Glazener, et al., 1990, Goldman, et al., 2014, Goverde, et al., 2000, Guzick, et al., 1999, Harrison and O'Moore, 1983, Ho, et al., 1998, Hughes, et al., 2004, Janko, et al., 1998, Karlstrom, et al., 1993, Kirby, et al., 1991, Leanza, et al., 2014, Martinez, et al., 1990, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006).

Included studies

Study design and setting

Of the 27 RCTs reporting on 4349 couples included in this systematic review, 21 had a parallel design (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Elzeiny, et al., 2014, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Goldman, et al., 2014, Goverde, et al., 2000, Guzick, et al., 1999, Ho, et al., 1998, Hughes, et al., 2004, Janko, et al., 1998, Karlstrom, et al., 1993, Kirby, et al., 1991, Leanza, et al., 2014, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006), and the other six were cross-over studies (Arici, et al., 1994, Crosignani, et al., 1991, Deaton, et al., 1990, Glazener, et al., 1990, Harrison and O'Moore, 1983, Martinez, et al., 1990). These studies were conducted in a variety of countries, including Netherlands (n = 5) (Bensdorp, et al., 2015, Custers, et al., 2011, Goverde, et al., 2000, Martinez, et al., 1990, Steures, et al., 2006), USA (n = 4) (Arici, et al., 1994, Deaton, et al., 1990, Goldman, et al., 2014, Guzick, et al., 1999), Italy (n = 3) (Arcaini, et al., 1996, Leanza, et al., 2014, Melis, et al., 1995), UK (n = 3)

(Bhattacharya, et al., 2008, Glazener, et al., 1990, Nandi, et al., 2017), Australia (n = 2) (Elzeiny, et al., 2014, Kirby, et al., 1991), Canada (n = 2) (Fisch, et al., 1989, Hughes, et al., 2004), India (n = 2) (Agarwal and Mittal, 2004, George, et al., 2006), China (n = 1) (Ho, et al., 1998), New Zealand (n = 1) (Farquhar, et al., 2017), Ireland (n = 1) (Harrison and O'Moore, 1983), Sweden (n = 1) (Karlstrom, et al., 1993), and Slovakia (n = 1) (Janko, et al., 1998). One study was conducted in a multi-country setting in Europe (Crosignani, et al., 1991).

Participants

These studies included 4349 couples with unexplained infertility. The mean female age across included studies ranged from 32 to 37 years, with most studies reporting a mean age younger than 35 years. The median or mean duration of infertility across included studies ranged from 23 to 78 months.

Interventions

One four-arm RCT compared expectant management, OS, IUI, and OS-IUI (Martinez, et al., 1990). We identified three three-arm RCTs: one compared expectant management, OS, and IUI (Bhattacharya, et al., 2008); another compared OS, OS-IUI, and IVF/ICSI (Crosignani, et al., 1991); and the third compared IUI, OS-IUI, and IVF/ICSI (Goverde, et al., 2000). The other 23 studies were two-arm studies. These studies compared OS versus expectant management (Fisch, et al., 1989, George, et al., 2006, Glazener, et al., 1990, Harrison and O'Moore, 1983), IUI versus expectant management (Kirby, et al., 1991), OS-IUI versus expectant management (Deaton, et al., 1990, Farquhar, et al., 2017, Steures, et al., 2006), IVF/ICSI versus expectant management (Hughes, et al., 2004), OS-IUI versus OS (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Ho, et al., 1998, Janko, et al., 1998, Karlstrom, et al., 1993, Melis, et al., 1995), OS-IUI versus IUI (Arici, et al., 1994, Guzick, et al., 1999, Leanza, et al., 2014), and IVF/ICSI versus OS-IUI (Bensdorp, et al., 2015, Custers, et al., 2011, Elzeiny, et al., 2014, Goldman, et al., 2014, Nandi, et al., 2017).

For RCTs comparing OS-IUI, IUI, and OS versus expectant management or each other, all compared the same number of cycles of different interventions - one cycle in five RCTs (Arici, et al., 1994, Crosignani, et al., 1991, Karlstrom, et al., 1993, Kirby, et al., 1991, Martinez, et al., 1990), three cycles in seven RCTs (Farquhar, et al., 2017, George, et al., 2006, Glazener, et al., 1990, Ho, et al., 1998, Janko, et al., 1998, Leanza, et al., 2014, Melis, et al., 1995), four cycles in three RCTs (Deaton, et al., 1990, Fisch, et al., 1989, Guzick, et al., 1999), five cycles in one RCT (Arcaini, et al., 1996), and six cycles in five RCTs (Agarwal and Mittal, 2004, Bhattacharya, et al., 2008, Goverde, et al., 2000, Harrison and O'Moore, 1983, Steures, et al., 2006).

For RCTs comparing IVF/ICSI with other interventions, (Hughes, et al., 2004) compared one cycle of IVF/ICSI versus three cycles of expectant management within 90 days; (Bensdorp, et al., 2015) compared three cycles of IVF/ICSI versus six cycles of

OS-IUI within 12 months; (Custers, et al., 2011) compared one cycle of IVF/ICSI versus three cycles of OS-IUI within four months; and (Nandi, et al., 2017) compared one cycle of IVF/ICSI versus three cycles of OS-IUI within six months. The other RCTs compared the same number of cycles of IVF versus other interventions without time limits: (Crosignani, et al., 1991) compared one cycle of IVF/ICSI with one cycle of OS and OS-IUI; (Elzeiny, et al., 2014) compared one cycle of IVF/ICSI versus one cycle of OS-IUI; (Goldman, et al., 2014) compared two cycles of IVF/ICSI versus two cycles of OS-IUI; and (Goverde, et al., 2000) compared six cycles of IVF/ICSI, six cycles of OS-IUI, and six cycles of IUI.

Elective or compulsive single embryo transfer policy was applied in three RCTs (Bensdorp, et al., 2015, Custers, et al., 2011, Nandi, et al., 2017). ICSI was used in three RCTs, only for couples with fertilisation failure in previous IVF or unexpected low sperm count on the day of oocyte retrieval (Bensdorp, et al., 2015, Goldman, et al., 2014, Nandi, et al., 2017).

Outcomes

Thirteen RCTs reported live birth (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Elzeiny, et al., 2014, Farquhar, et al., 2017, George, et al., 2006, Goldman, et al., 2014, Goverde, et al., 2000, Guzick, et al., 1999, Hughes, et al., 2004, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006), and 14 RCTs reported multiple pregnancy (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Deaton, et al., 1990, Elzeiny, et al., 2014, Farquhar, et al., 2017, George, et al., 2006, Glazener, et al., 1990, Goldman, et al., 2014, Goverde, et al., 2000, Ho, et al., 1998, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). Twenty-six studies reported clinical pregnancy (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Arici, et al., 1994, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Crosignani, et al., 1991, Custers, et al., 2011, Deaton, et al., 1990, Elzeiny, et al., 2014, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Glazener, et al., 1990, Goldman, et al., 2014, Guzick, et al., 1999, Harrison and O'Moore, 1983, Ho, et al., 1998, Hughes, et al., 2004, Janko, et al., 1998, Karlstrom, et al., 1993, Kirby, et al., 1991, Leanza, et al., 2014, Martinez, et al., 1990, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). Eight studies reported moderate/severe OHSS as an outcome (Bensdorp, et al., 2015, Deaton, et al., 1990, Elzeiny, et al., 2014, Goldman, et al., 2014, Goverde, et al., 2000, Ho, et al., 1998, Melis, et al., 1995, Nandi, et al., 2017).

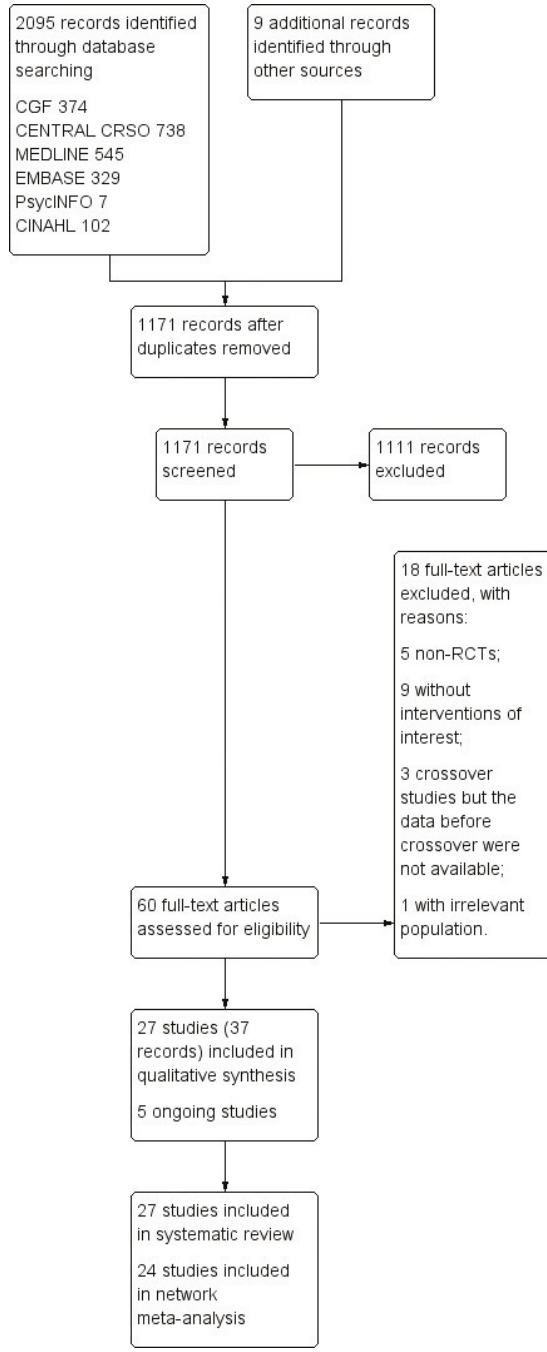


Figure 1. Study flow diagram.

Excluded studies

We excluded 18 studies from the review for the following reasons (Figure 1): five were non-RCTs (Fujii, et al., 1997, Nulsen, et al., 1993, Prentice, et al., 1995, Tjon-Kon-Fat, et al., 2014, Zayed, et al., 1997); nine did not include interventions of interest (Buvat, et al., 1993, Chung, et al., 1995, Goldman, et al., 2010, Leanza, et al., 2014, Melis, et al., 1987, Murdoch, et al., 1991, Reindollar, et al., 2010, Shokeir, 2006, Soliman, et al., 1993); three were cross-over studies but the data before cross-over were not available (Gregoriou, et al., 1995, Martinez, et al., 1991, Zikopoulos, et al., 1993); and one had an irrelevant population (i.e. included women with polycystic ovary syndrome) (Zolghadri, et al., 2012).

We identified five ongoing studies from Belgium (NCT01992731, 2013), China (NCT03455426, 2018), Egypt (NCT02461173, 2015), France (NCT02001870, 2013), and Netherlands (NTR5599, 2016), respectively.

Risk of bias in included studies

Allocation (selection bias)

Sequence generation

As shown in Figure 2 and Figure 3, 12 studies reported adequate methods for random sequence generation and therefore were rated as low risk of bias in sequence generation (Agarwal and Mittal, 2004, Arici, et al., 1994, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Elzeiny, et al., 2014, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Goverde, et al., 2000, Nandi, et al., 2017, Steures, et al., 2006). The other 16 studies did not describe the method used and were rated as unclear risk for this domain.

Allocation concealment

Twelve studies described adequate methods for allocation concealment (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Elzeiny, et al., 2014, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Goldman, et al., 2014, Goverde, et al., 2000, Hughes, et al., 2004, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006), and the other 16 studies did not describe methods of allocation concealment and were scored as unclear risk of bias for this domain.

Blinding (performance bias and detection bias)

Blinding of participants and personnel (performance bias)

Five studies were rated as low risk of performance bias as placebos were used (Fisch, et al., 1989, George, et al., 2006, Glazener, et al., 1990, Harrison and O'Moore, 1983, Leanza, et al., 2014). The remaining studies were rated as high risk of performance bias as they were not blinded, although blinding was not possible due to the nature of the interventions.

Blinding of outcome assessors (detection bias)

Given that our outcomes of interest were objective outcomes, we considered that blinding was unlikely to impact these outcomes. Therefore, all studies were rated as low risk of detection bias.

Incomplete outcome data (attrition bias)

Three studies had 19%, 20%, and 21% incomplete outcome data, respectively, and therefore were rated as high risk of attrition bias (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Deaton, et al., 1990). Thirteen studies had low risk of attrition bias (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Farquhar, et al., 2017, Glazener, et al., 1990, Goldman, et al., 2014, Guzick, et al., 1999, Harrison and O'Moore, 1983, Hughes, et al., 2004, Martinez, et al., 1990, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006) and the other 11 studies were scored as unclear risk.

Selective reporting (reporting bias)

Two studies did not report the outcome data for each group separately and were rated as high risk of reporting bias (Agarwal and Mittal, 2004, Arcaini, et al., 1996). Twelve studies reported both live birth and multiple pregnancy and were rated as low risk of reporting bias (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Elzeiny, et al., 2014, Farquhar, et al., 2017, George, et al., 2006, Goldman, et al., 2014, Goverde, et al., 2000, Hughes, et al., 2004, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). The other 14 studies were scored as unclear risk.

Other potential sources of bias

There was disagreement on the number of participants in the methods and results sections in one study and this was rated as high risk of bias (Glazener, et al., 1990). Thirteen studies were scored as low risk of other bias (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Elzeiny, et al., 2014, Farquhar, et al., 2017, Goldman, et al., 2014, Goverde, et al., 2000, Guzick, et al., 1999, Hughes, et al., 2004, Nandi, et al., 2017, Steures, et al., 2006). The other 14 studies were scored as unclear risk.

Figure 2. Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

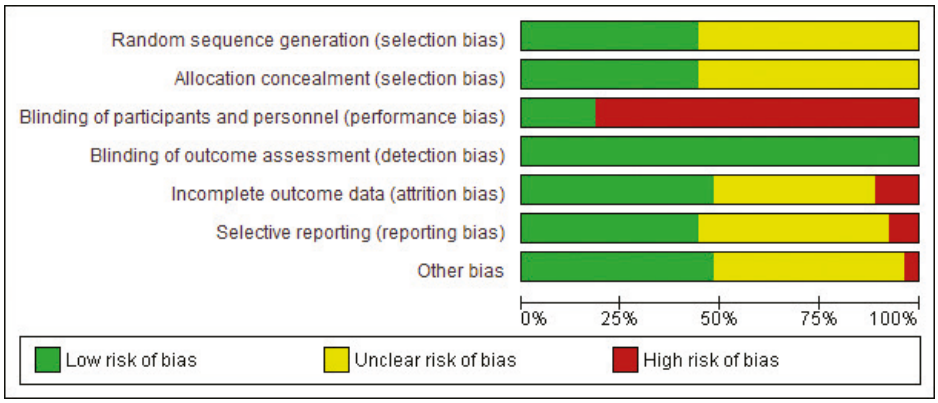
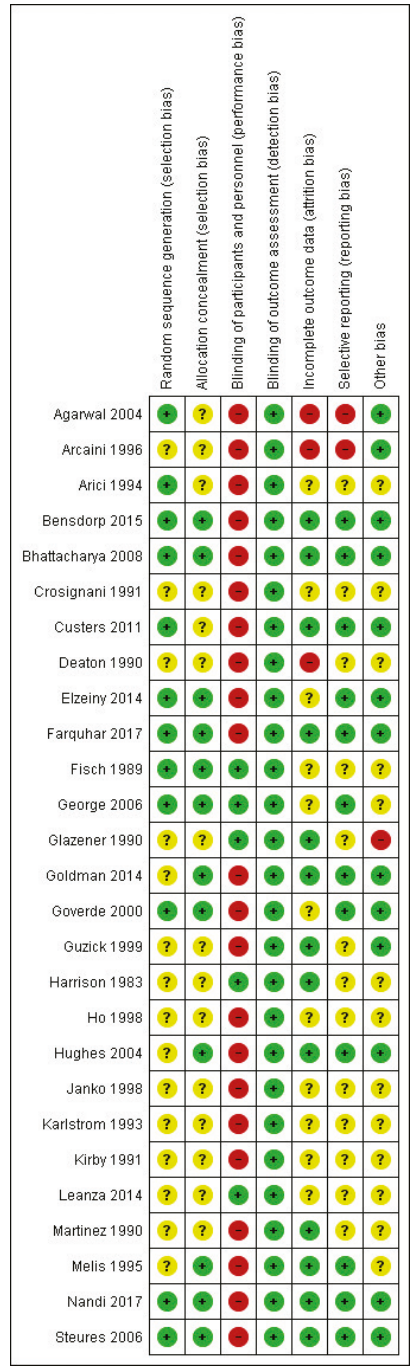


Figure 3. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Effects of interventions

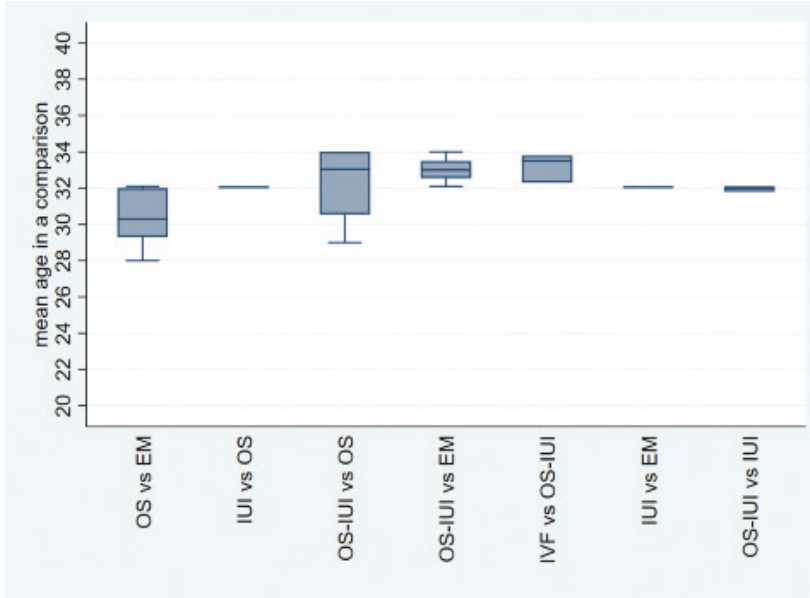
Network meta-analysis

Based on above-mentioned Unit of analysis issues, two RCTs (Elzeiny, et al., 2014, Goldman, et al., 2014) and IVF/ICSI arms in two other RCTs (Crosignani, et al., 1991, Goverde, et al., 2000) were excluded from this network meta-analysis, as these RCTs compared IVF/ICSI and other interventions in the same number of cycles. We further excluded (Hughes, et al., 2004) from this network meta-analysis, as this RCT allowed transfer of up to four embryos. The remaining RCTs comparing IVF/ICSI all used single embryo transfer policy. Detailed data analyses for these five RCTs that were excluded from the network meta-analysis are presented in Analysis 3.1, Analysis 3.2, and Analysis 3.3. Finally, 24 RCTs reporting on 3983 couples with unexplained infertility were included in this network meta-analysis.

We observed high heterogeneity in the pairwise meta-analysis of OS-IUI and expectant management (EM) (I^2 = 91% for live birth). This is likely due to clinical heterogeneity among participants in the two included RCTs - (Steures, et al., 2006) included couples with intermediate prognosis of natural conception, and (Farquhar, et al., 2017) included couples with poor prognosis of natural conception. Both RCTs applied an existing prediction model to estimate the prognosis of natural conception (Hunault, et al., 2004). We included these RCTs in this network meta-analysis to estimate the average treatment effect in this comparison, and we downgraded the certainty of evidence due to heterogeneity based on criteria described in the methods. To further assess robustness of the evidence, we performed two additional post-hoc sensitivity analyses: excluding expectant management from the network; and limiting to RCTs including couples with poor prognosis of natural conception.

We assessed the transitivity assumption in this network meta-analysis by evaluating two potential effect modifiers: age and duration of infertility. The distribution of mean age in different studies across different comparisons is presented in Figure 4. The median value of mean age across different comparisons is around 32 years. Duration of infertility is very unlikely to be normally distributed; therefore reporting the mean seems inappropriate and can lead to overestimation of the median value. However, 10 RCTs reported mean duration of infertility (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Arici, et al., 1994, Deaton, et al., 1990, Fisch, et al., 1989, Goverde, et al., 2000, Guzick, et al., 1999, Harrison and O'Moore, 1983, Martinez, et al., 1990, Melis, et al., 1995), and seven other RCTs did not report median or mean duration of infertility (Crosignani, et al., 1991, George, et al., 2006, Ho, et al., 1998, Janko, et al., 1998, Karlstrom, et al., 1993, Kirby, et al., 1991, Leanza, et al., 2014). Therefore, it is impossible for us to assess the distribution of duration of infertility across different comparisons. However, as these five interventions are jointly randomisable for any participant with unexplained infertility, we considered the transitivity assumption valid.

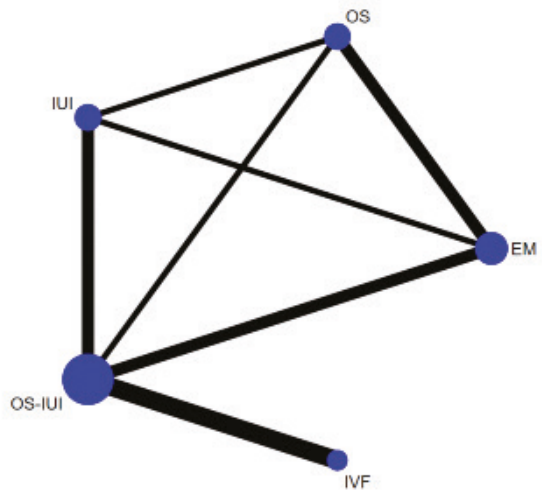
Figure 4. Box plot for the distribution of means of age in different studies across different comparisons.



Live birth

Ten studies reported live birth (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Farquhar, et al., 2017, George, et al., 2006, Goverde, et al., 2000, Guzick, et al., 1999, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). These RCTs included 2725 couples with unexplained infertility. A network plot for live birth is presented in Figure 5. Three RCTs compared IVF/ICSI versus OS-IUI (Bensdorp, et al., 2015, Custers, et al., 2011, Nandi, et al., 2017); two RCTs compared OS-IUI versus IUI (Goverde, et al., 2000, Guzick, et al., 1999); two RCTs compared OS versus expectant management (Bhattacharya, et al., 2008, George, et al., 2006); two RCTs compared OS-IUI versus expectant management (Farquhar, et al., 2017, Steures, et al., 2006); one RCT compared IUI versus expectant management (Bhattacharya, et al., 2008); and one RCT compared OS-IUI versus OS (Melis, et al., 1995).

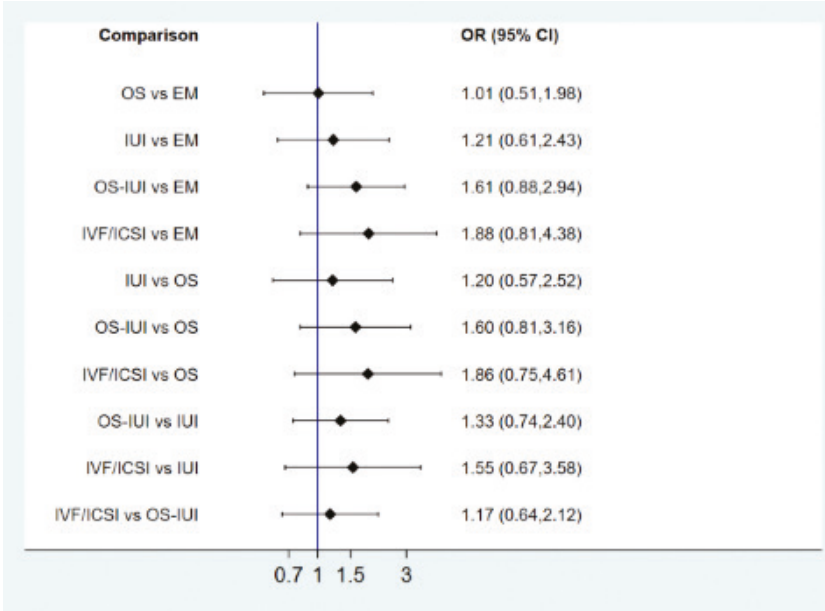
Figure 5. Network plot for live birth.



Each node represents an intervention, and the size of each node is proportional to the number of trials reporting such intervention. The widths of the lines are proportional to the numbers of trials comparing each pair of interventions.

The results of the network meta-analysis are shown in Figure 6. They showed insufficient evidence of a difference between OS, IUI, OS-IUI, or IVF/ICSI and expectant management (odds ratio (OR) 1.01, 95% confidence interval (CI) 0.51 to 1.98; low-certainty evidence; OR 1.21, 95% CI 0.61 to 2.43; low-certainty evidence; OR 1.61, 95% CI 0.88 to 2.94; low-certainty evidence; OR 1.88, 95% CI 0.81 to 4.38; low-certainty evidence). These data suggest that if the chance of live birth following expectant management is assumed to be 16.6%, the chance following OS, IUI, OS-IUI, and IVF would be 9.2% to 28.2%, 10.8% to 32.5%, 14.9% to 36.9%, and 13.9% to 46.5%, respectively.

Figure 6. Network meta-analysis for live birth.



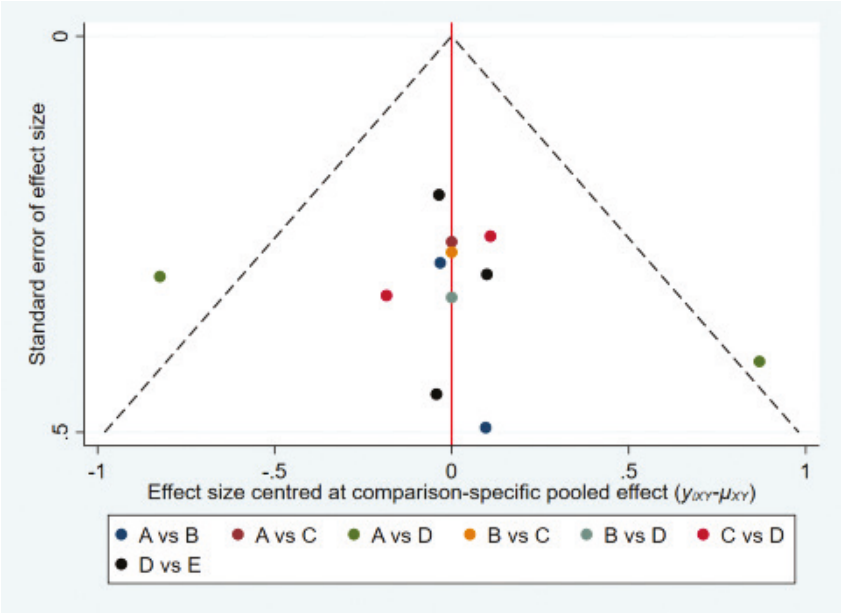
Each diamond represents the estimate summary odds ratio of each comparison; each horizontal line represents the confidence interval of each comparison; blue vertical line represents line of no effect (odds ratio = 1). Odds ratio greater than 1 favours the first intervention; odds ratio less than 1 favours the second intervention.

Evidence of a difference between IUI and OS (OR 1.20, 95% CI 0.57 to 2.52; low-certainty evidence), OS-IUI and OS (OR 1.60, 95% CI 0.81 to 3.16; low-certainty evidence), IVF/ICSI and OS (OR 1.86, 95% CI 0.75 to 4.61; low-certainty evidence), OS-IUI and IUI (OR 1.33, 95% CI 0.74 to 2.40; low-certainty evidence), IVF/ICSI and IUI (OR 1.55, 95% CI 0.67 to 3.58; low-certainty evidence), or IVF/ICSI and OS-IUI (OR 1.17, 95% CI 0.64 to 2.12; low-certainty evidence) was insufficient. Overall certainty of evidence in all comparisons was low due to concerns regarding imprecision and heterogeneity.

Results show no evidence of global inconsistency (P = 0.55) or local inconsistency in the network meta-analysis on live birth. The comparison-adjusted funnel plot seems symmetrical, implying the absence of small study effects in this network (Figure 7). Cumulative rankograms illustrate the probability per rank for each treatment in terms of live birth (Figure 8). The SUCRA values for expectant management, OS, IUI, OS-IUI, and IVF/ICSI were 23.1%, 24.1%, 43.7%, 74.2%, and 85.0%, respectively. This suggests that among all interventions, IVF/ICSI is more likely to result in more live births than the other interventions, followed by OS-IUI, IUI, OS, and expectant management.

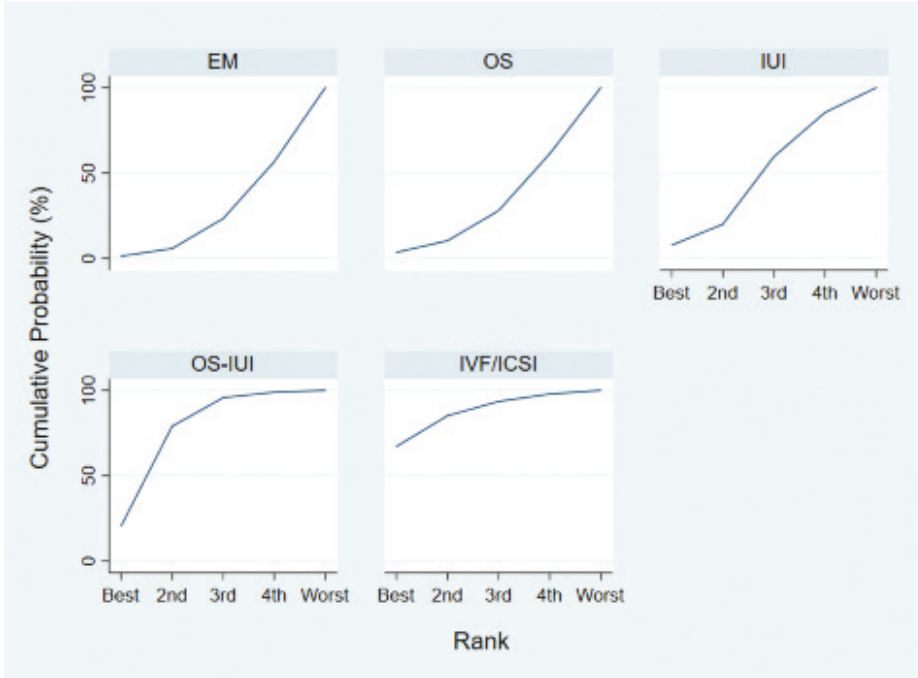
Results of pairwise meta-analyses are presented in Analysis 1.1. Overall, results were consistent with those in network meta-analysis. As most comparisons included a very limited number of studies, wide confidence intervals were observed in all comparisons, implying imprecision of the evidence.

Figure 7. Comparison-adjusted funnel plot for live birth.



(A: expectant management; B: OS; C: IUI; D: OS-IUI; E: IVF/ICSI.)

Figure 8. Cumulative rankograms of interventions for live birth.



Each cumulative rankogram illustrates the cumulative probability of each ranking (from the best to the worst rank) for each intervention in terms of live birth.

Subgroup analyses

Women ≤ 38 years versus women > 38 years

One RCT did not report details of age in the inclusion criteria or results (George, et al., 2006), and the other RCTs all reported a mean age < 35 years. As the breakdown data for women in different age groups were not available, this subgroup analysis was not performed.

Short duration of infertility (≤ 2 years) versus long duration of infertility (> 2 years)

As the breakdown data for women in different age groups were not available, we used median duration of infertility in different RCTs for this subgroup analysis. Therefore this subgroup analysis should be interpreted with caution, given that it was not based on the breakdown data for different groups.

One study did not report details of the duration of infertility in the inclusion criteria or the results (George, et al., 2006); therefore we excluded this study from the subgroup analysis. Two studies included couples with a median or mean duration of infertility ≤ 2 years (Nandi, et al., 2017, Steures, et al., 2006). One compared IVF/ICSI versus OS-IUI

(Nandi, et al., 2017), and the other compared IVF/ICSI versus expectant management (Steures, et al., 2006). Network meta-analysis is presented in Figure 9. Evidence of a difference in live birth between OS-IUI or IVF/ICSI and expectant management was insufficient (OR 0.82, 95% CI 0.45 to 1.49; OR 1.05, 95% CI 0.46 to 2.43). Seven studies reported median duration of infertility > 2 years (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Farquhar, et al., 2017, Goverde, et al., 2000, Guzick, et al., 1999, Melis, et al., 1995). Network meta-analysis of these studies is presented in Figure 10. Effect sizes of IVF/ICSI and OS-IUI versus expectant management were larger than those in the main analysis.

IVF/ICSI with single embryo transfer policy and IVF/ICSI with non-single embryo transfer policy

As all RCTs including an IVF/ICSI arm applied single embryo transfer policy, this subgroup analysis was not performed.

Figure 9. Subgroup analysis for live birth - RCTs with a median duration of infertility ≤ 2 years.

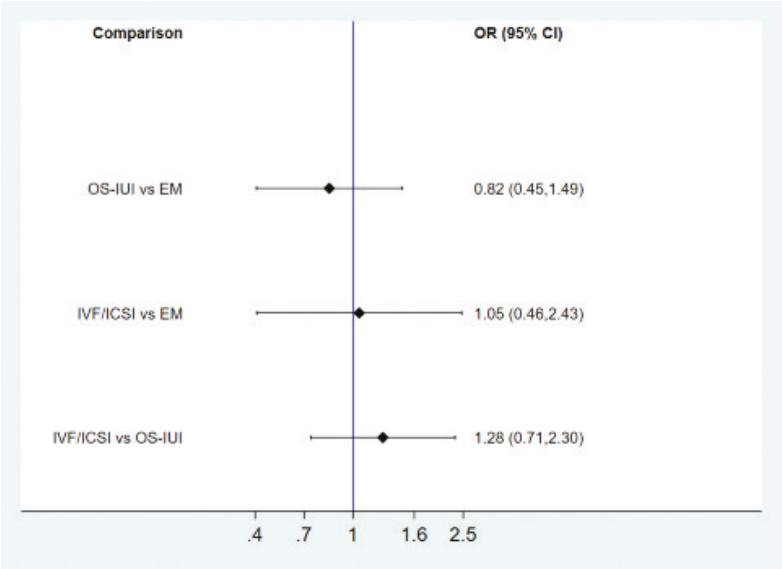
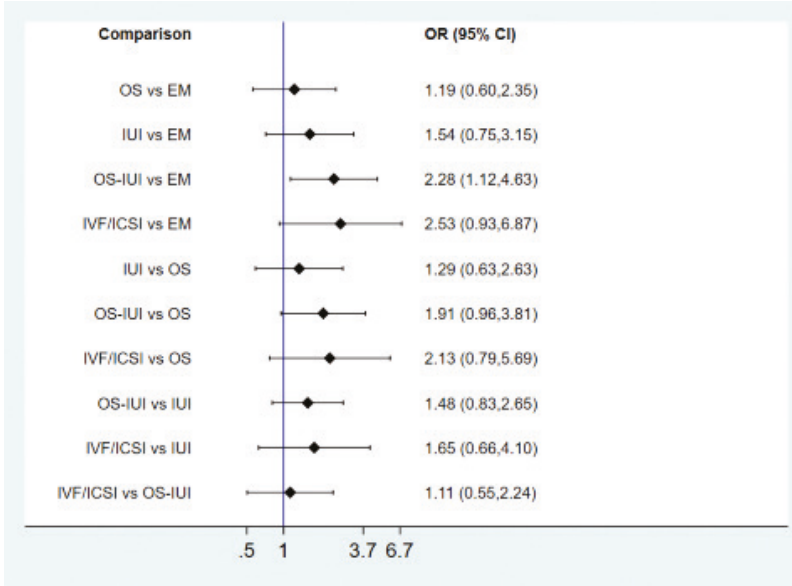


Figure 10. Subgroup analysis for live birth - RCTs with a median duration of infertility >2 years.



Sensitivity analyses

Restricting to RCTs with no domains at high risk of bias
Most RCTs were rated at high risk of performance bias; therefore this analysis was not possible.

Excluding participants with missing outcome data
After participants with missing outcome data were excluded, the results of network meta-analysis were consistent with the main analysis in all comparisons (Figure 11).

Excluding abstract-only publications
One abstract was excluded from this sensitivity analysis (George, et al., 2006). Results of this sensitivity analysis were consistent with those of the main analysis for all comparisons (Figure 12).

Including only RCTs with the outcome live birth
All 10 studies reported live birth; therefore this analysis was not performed.

Excluding expectant management from the network
Results of network meta-analysis of the remaining four interventions were consistent with results of the main analysis (Figure 13).

Restricting to RCTs including couples with poor prognosis of natural conception
Three RCTs (Bensdorp, et al., 2015, Custers, et al., 2011, Farquhar, et al., 2017) included couples with poor prognosis of natural conception based on an existing prediction

model (Hunault, et al., 2004). Network meta-analysis (Figure 14) showed that compared to expectant management, OS-IUI (OR 4.48, 95% CI 2.00 to 10.1; moderate-certainty evidence) or IVF/ICSI (OR 4.99, 95 CI 2.07 to 12.04; moderate-certainty evidence) increased the odds of live birth, and there was insufficient evidence of a difference between IVF/ICSI and OS-IUI (OR 1.11, 95% CI 0.78 to 1.60; low-certainty evidence). This sensitivity analysis showed the clinically important differences of OS-IUI and IVF/ICSI versus expectant management.

Figure 11. Sensitivity analysis for live birth by exclusion of participants with missing outcome data.

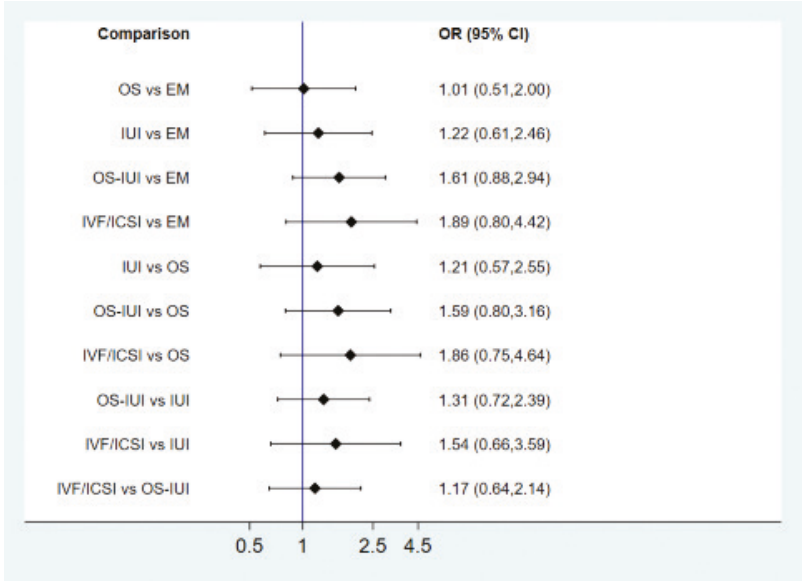


Figure 12. Sensitivity analysis for live birth by exclusion of abstract-only publications.

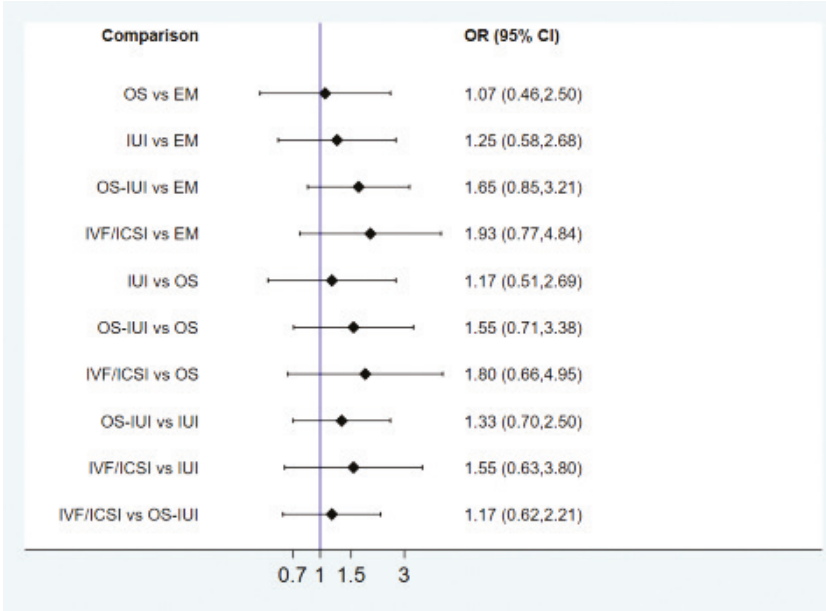


Figure 13. Sensitivity analysis for live birth by excluding RCTs involving expectant management from the network

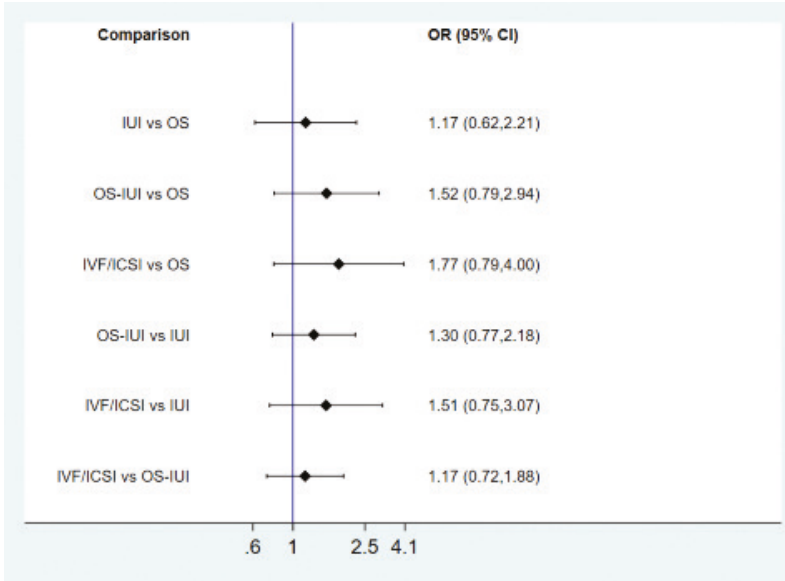
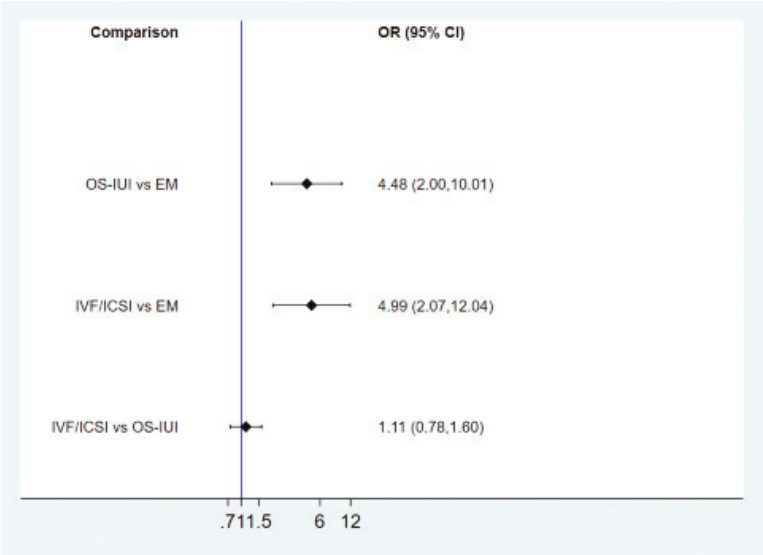


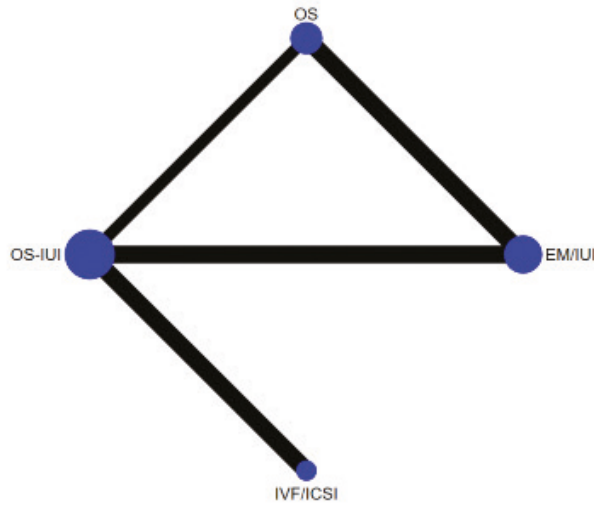
Figure 14. Sensitivity analysis for live birth by limiting to RCTs on couples with poor prognosis of natural conception.



Multiple pregnancy

One study reported 0 events in both groups and was excluded from the analysis (Deaton, et al., 1990). Eleven RCTs reporting on 2564 couples were included in the network meta-analysis of multiple pregnancy (Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Custers, et al., 2011, Farquhar, et al., 2017, George, et al., 2006, Glazener, et al., 1990, Goverde, et al., 2000, Ho, et al., 1998, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). The network plot for multiple pregnancy is presented in Figure 15.

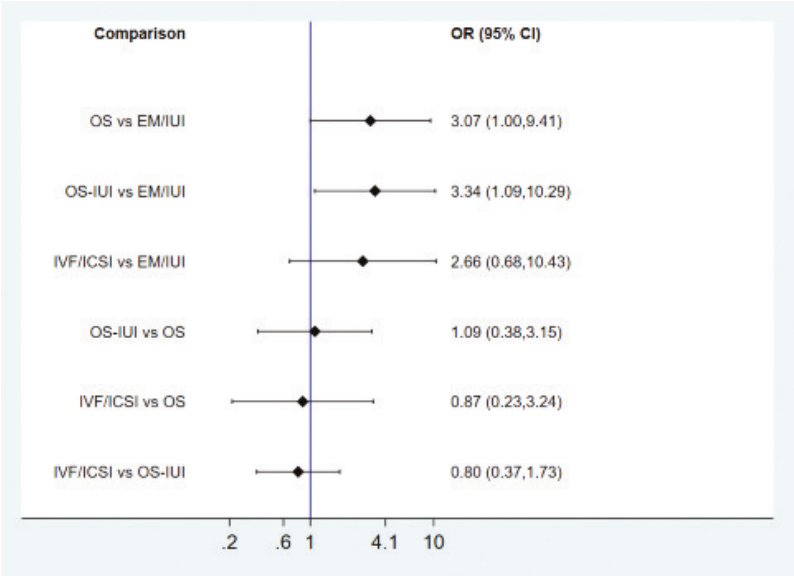
Figure 15. Network plot for multiple pregnancy.



Results of network meta-analysis are shown in Figure 16. Compared to expectant management/IUI, OS (OR 3.07, 95% CI 1.00 to 9.41; low-certainty evidence) or OS-IUI (OR 3.34, 95% CI 1.09 to 10.29; moderate-certainty evidence) increased the odds of multiple pregnancy, and there was insufficient evidence of a difference between IVF/ICSI and expectant management/IUI (OR 2.66, 95% CI 0.68 to 10.43; low-certainty evidence). These findings suggest that if the chance of multiple pregnancy following expectant management or IUI is assumed to be 0.6%, the chance following OS, OS-IUI, and IVF/ICSI would be 0.6% to 5.0%, 0.6% to 5.4%, and 0.4% to 5.5%, respectively.

These was insufficient evidence of a difference between OS-IUI and OS (OR 1.09, 95% CI 0.38 to 3.15; very-low-certainty evidence), IVF/ICSI and OS (OR 0.87, 95% CI 0.23 to 3.24; low-certainty evidence), or IVF/ICSI and OS-IUI (OR 0.80, 95% CI 0.37 to 1.73; low-certainty evidence).

Figure 16. Network meta-analysis for multiple pregnancy.



There was no evidence of global inconsistency (P = 0.34) or local inconsistency in the network meta-analysis on multiple pregnancy. Cumulative rankograms illustrate the probability per rank for each treatment in terms of multiple pregnancy (Figure 17). The comparison-adjusted funnel plot seems symmetrical, implying the absence of small study effects in this network (Figure 18). The SUCRA values for expectant management/IUI, OS, OS-IUI, and IVF/ICSI were 95.3%, 33.8%, 24.5%, and 46.4%, respectively. This suggests that expectant management/IUI was more likely to result in fewer multiple pregnancies than other interventions, followed by IVF/ICSI, OS, and OS-IUI.

Results of pairwise meta-analyses (Analysis 1.2) are consistent with those in the network meta-analysis.

Figure 17. Cumulative rankograms of interventions for multiple pregnancy.

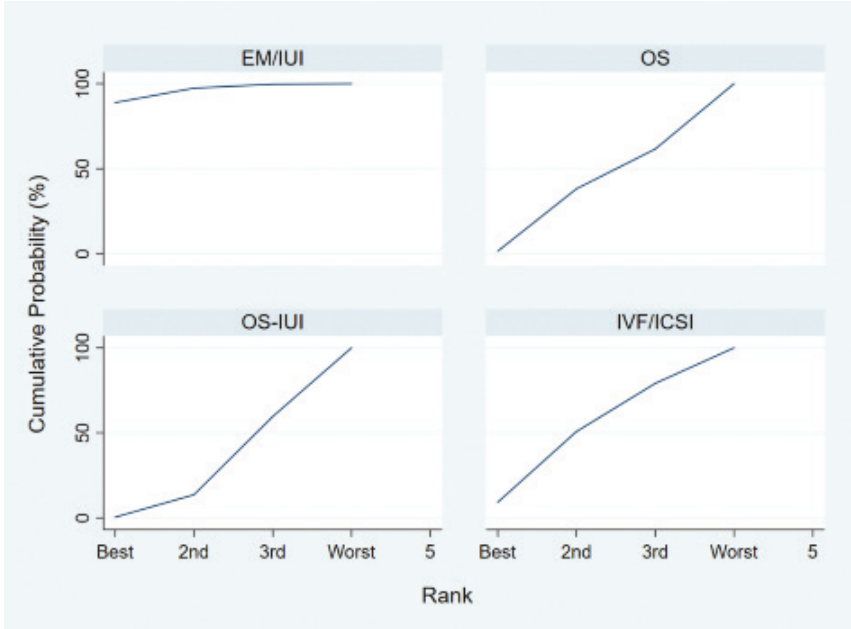
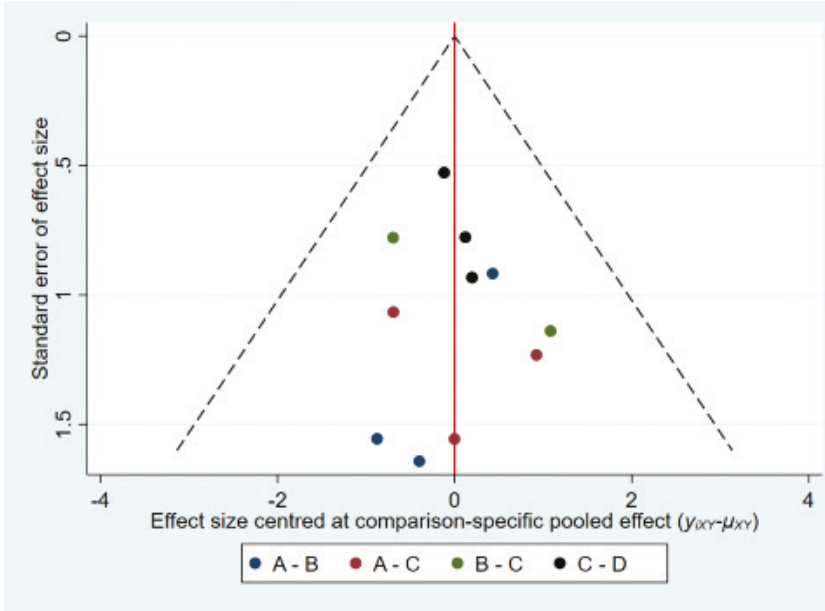


Figure 18. Comparison-adjusted funnel plot for multiple pregnancy.

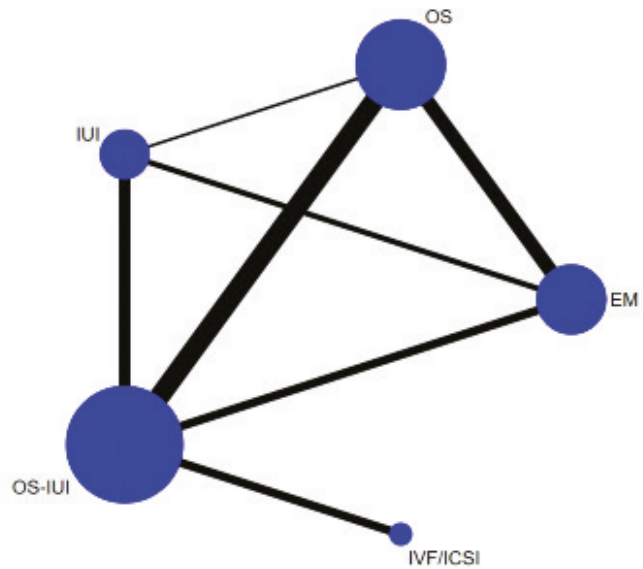


(A: expectant management or IUI; B: OS; C: OS-IUI; D: IVF/ICSI.)

Clinical pregnancy

Twenty-three RCTs reporting on 3792 couples were included in the network meta-analysis of clinical pregnancy (Agarwal and Mittal, 2004, Arcaini, et al., 1996, Arici, et al., 1994, Bensdorp, et al., 2015, Bhattacharya, et al., 2008, Crosignani, et al., 1991, Custers, et al., 2011, Deaton, et al., 1990, Farquhar, et al., 2017, Fisch, et al., 1989, George, et al., 2006, Glazener, et al., 1990, Guzick, et al., 1999, Harrison and O'Moore, 1983, Ho, et al., 1998, Janko, et al., 1998, Karlstrom, et al., 1993, Kirby, et al., 1991, Leanza, et al., 2014, Martinez, et al., 1990, Melis, et al., 1995, Nandi, et al., 2017, Steures, et al., 2006). The network plot for clinical pregnancy is presented in Figure 19.

Figure 19. Network plot for clinical pregnancy.



Results of the network meta-analysis are shown in Figure 20. Compared to expectant management, OS-IUI or IVF/ICSI increased the odds of live birth (OR 2.32, 95% CI 1.39 to 3.90; low-certainty evidence; OR 3.03, 95% CI 1.32 to 6.94; low-certainty evidence). There was insufficient evidence of a difference between OS and expectant management (OR 1.64, 95% CI 0.99 to 2.73; very-low-certainty evidence) or between IUI and expectant management (OR 1.20, 95% CI 0.61 to 2.36; low-certainty evidence). These findings suggest that if the chance of clinical pregnancy following expectant management is assumed to be 16.4%, the chance following OS, IUI, OS-IUI, and IVF/

ICSI would be 15.5% to 33.7%, 10.2% to 30.5%, 20.5% to 42.0%, and 19.7% to 56.3%, respectively.

Compared to OS, IVF/ICSI increased the odds of clinical pregnancy (OR 1.84, 95% CI 1.40 to 4.02; low-certainty evidence). There was insufficient evidence of a difference between IUI or OS-IUI and expectant management (OR 0.73, 95% CI 0.38 to 1.42; very low-certainty evidence; OR 1.41, 95% CI 0.92 to 2.18; very low-certainty evidence). Compared to IUI, OS-IUI or IVF/ICSI increased the odds of clinical pregnancy (OR 1.94, 95% CI 1.05 to 3.57; very low-certainty evidence; OR 2.52, 95% CI 1.04 to 6.16; low-certainty evidence). Evidence of a difference between IVF/ICSI and OS-IUI for clinical pregnancy was insufficient (OR 1.30, 95% CI 0.68 to 2.50; low-certainty evidence).

There was no evidence of global inconsistency (P = 0.23), but local inconsistency was detected in the comparison between IUI and OS (P = 0.039). Therefore, the certainty of evidence in this comparison was downgraded due to incoherence. Cumulative rankograms illustrate the cumulative probability per rank for each treatment in terms of clinical pregnancy (Figure 21). The comparison-adjusted funnel plot seems symmetrical, implying the absence of small study effects in this network (Figure 22). The SUCRA values for expectant management, OS, IUI, OS-IUI, and IVF/ICSI were 7.8%, 48.4%, 23.3%, 78.8%, and 91.7%, respectively. This suggests that IVF/ICSI was is more likely to result in more clinical pregnancies than the other interventions, followed by OS-IUI, OS, IUI, and expectant management.

Results of pairwise meta-analyses were consistent with those in the network meta-analysis (Analysis 1.3).

Figure 20. Network meta-analysis for clinical pregnancy.

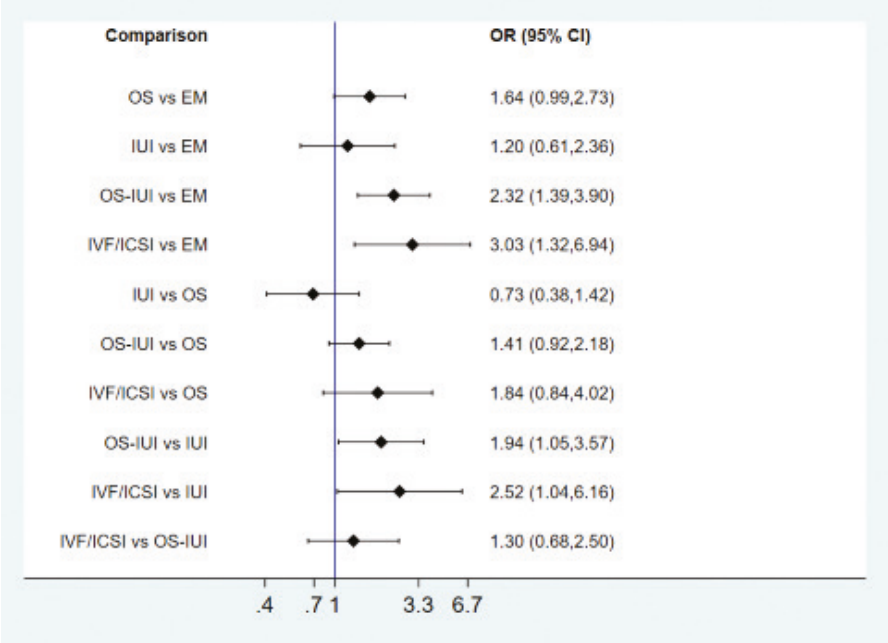


Figure 21. Cumulative rankograms of interventions for clinical pregnancy.

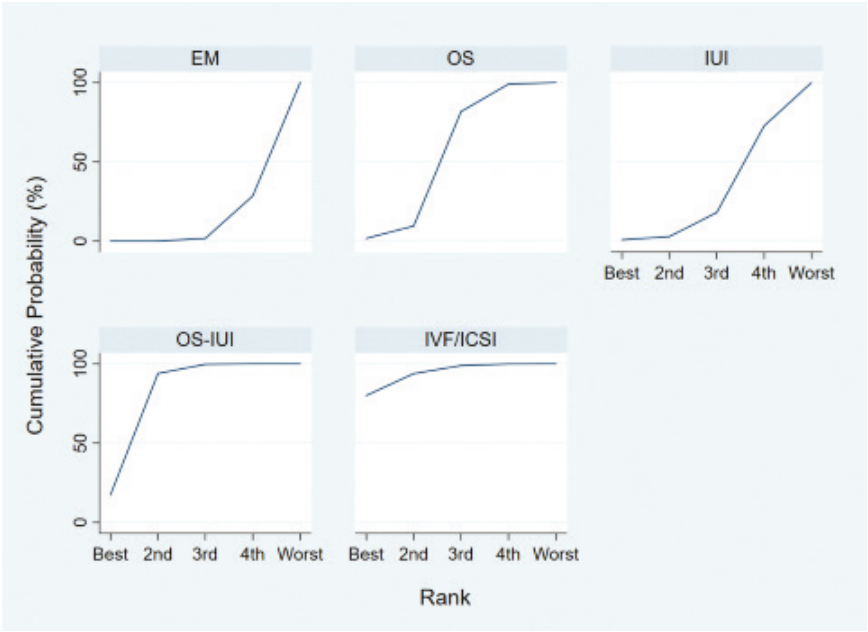
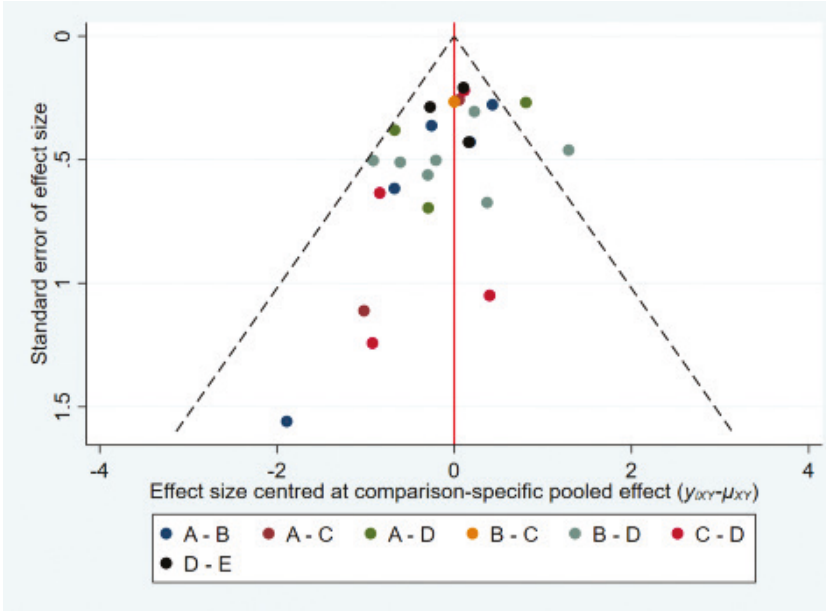


Figure 22. Comparison-adjusted funnel plot for clinical pregnancy.



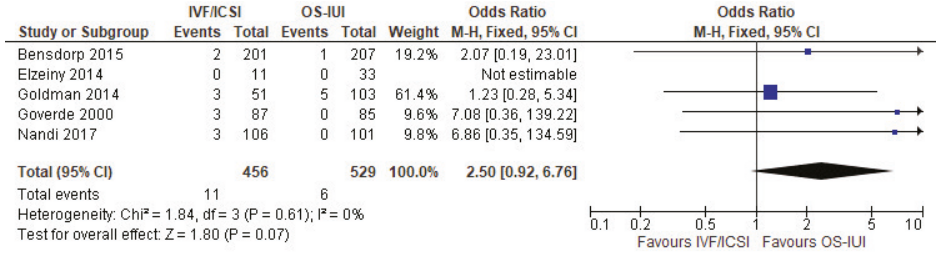
(A: expectant management; B: OS; C: IUI; D: OS-IUI; E: IVF/ICSI.)

OHSS

Eight studies reported moderate/severe OHSS. Four studies reported zero events in both groups (Deaton, et al., 1990, Elzeiny, et al., 2014, Ho, et al., 1998, Melis, et al., 1995). We did not perform network meta-analysis given the extremely low event rates for some interventions.

Five studies compared IVF/ICSI versus OS-IUI (Bensdorp, et al., 2015, Elzeiny, et al., 2014, Goldman, et al., 2014, Goverde, et al., 2000, Nandi, et al., 2017). Pooled analysis showed insufficient evidence of a difference between IVF/ICSI and OS-IUI (OR 2.50, 95% CI 0.92 to 6.76; 5 studies; 985 women; moderate-certainty evidence; Figure 23). This suggests that if the chance of moderate/severe OHSS following OS-IUI is assumed to be 1.1%, the chance following IVF/ICSI would be between 1.0% and 7.2%.

Figure 23. Forest plot of comparison: 2 Pairwise meta-analysis for OHSS, outcome: 2.5 IVF/ICSI vs OS-IUI.



DATA AND ANALYSES

1. Pairwise meta-analyses for live birth, multiple pregnancy, and clinical pregnancy

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
1.1 Live birth	10		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
1.1.1 OS vs EM	2	527	Odds Ratio (M-H, Random, 95% CI)	0.80 [0.49, 1.31]
1.1.2 IUI vs EM	1	386	Odds Ratio (M-H, Random, 95% CI)	1.44 [0.87, 2.40]
1.1.3 OS-IUI vs EM	2	454	Odds Ratio (M-H, Random, 95% CI)	1.88 [0.36, 9.90]
1.1.4 IUI vs OS	1	387	Odds Ratio (M-H, Random, 95% CI)	1.85 [1.09, 3.16]
1.1.5 OS-IUI vs OS	1	184	Odds Ratio (M-H, Random, 95% CI)	0.88 [0.46, 1.67]
1.1.6 OS-IUI vs IUI	2	636	Odds Ratio (M-H, Random, 95% CI)	1.68 [1.14, 2.49]
1.1.7 IVF/ICSI vs OS-IUI	3	731	Odds Ratio (M-H, Random, 95% CI)	1.16 [0.85, 1.57]
1.2 Multiple pregnancy	12		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
1.2.1 OS vs EM/IUI	3	934	Odds Ratio (M-H, Random, 95% CI)	2.04 [0.51, 8.24]
1.2.2 OS-IUI vs EM/IUI	4	676	Odds Ratio (M-H, Random, 95% CI)	5.04 [1.24, 20.49]
1.2.3 OS-IUI vs OS	2	274	Odds Ratio (M-H, Random, 95% CI)	0.69 [0.12, 3.81]
1.2.5 IVF/ICSI vs OS-IUI	3	731	Odds Ratio (M-H, Random, 95% CI)	0.80 [0.37, 1.73]
1.3 Clinical pregnancy	23		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
1.3.1 OS vs EM	6	939	Odds Ratio (M-H, Random, 95% CI)	1.31 [0.82, 2.10]
1.3.2 IUI vs EM	3	528	Odds Ratio (M-H, Random, 95% CI)	1.52 [0.93, 2.47]
1.3.3 OS-IUI vs EM	4	525	Odds Ratio (M-H, Random, 95% CI)	2.69 [0.96, 7.55]
1.3.4 IUI vs OS	2	407	Odds Ratio (M-H, Random, 95% CI)	1.69 [1.01, 2.82]
1.3.5 OS-IUI vs OS	8	763	Odds Ratio (M-H, Random, 95% CI)	1.26 [0.73, 2.18]
1.3.6 OS-IUI vs IUI	4	579	Odds Ratio (M-H, Random, 95% CI)	2.56 [1.72, 3.80]
1.3.7 IVF/ICSI vs OS-IUI	3	731	Odds Ratio (M-H, Random, 95% CI)	1.29 [0.95, 1.76]

2. Pairwise meta-analysis for OHSS

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
2.1 OS-IUI vs EM	1	51	Odds Ratio (M-H, Fixed, 95% CI)	Not estimable
2.2 OS-IUI vs OS	2	274	Odds Ratio (M-H, Fixed, 95% CI)	Not estimable
2.3 OS-IUI vs IUI	1	171	Odds Ratio (M-H, Fixed, 95% CI)	Not estimable
2.4 IVF/ICSI vs IUI	1	173	Odds Ratio (M-H, Fixed, 95% CI)	7.17 [0.36, 140.84]
2.5 IVF/ICSI vs OS-IUI	5	985	Odds Ratio (M-H, Fixed, 95% CI)	2.50 [0.92, 6.76]

3. Data analyses of RCTs that were not included in the network meta-analysis

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
3.1 Live birth	4		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
3.1.1 IVF/ICSI vs EM	1	51	Odds Ratio (M-H, Random, 95% CI)	22.00 [2.56, 189.37]
3.1.2 IVF/ICSI vs IUI	1	173	Odds Ratio (M-H, Random, 95% CI)	1.49 [0.79, 2.82]
3.1.3 IVF/ICSI vs OS-IUI	3	370	Odds Ratio (M-H, Random, 95% CI)	2.23 [0.83, 5.98]
3.2 Multiple pregnancy	3		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
3.2.1 IVF/ICSI vs IUI	1	173	Odds Ratio (M-H, Random, 95% CI)	7.44 [0.90, 61.80]
3.2.2 IVF/ICSI vs OS-IUI	3	370	Odds Ratio (M-H, Random, 95% CI)	0.81 [0.37, 1.74]
3.3 Clinical pregnancy	4		Odds Ratio (M-H, Random, 95% CI)	Subtotals only
3.3.1 IVF/ICSI vs EM	1	51	Odds Ratio (M-H, Random, 95% CI)	8.00 [1.89, 33.85]
3.3.2 IVF/ICSI vs OS	1	103	Odds Ratio (M-H, Random, 95% CI)	2.36 [0.72, 7.72]
3.3.3 IVF/ICSI vs OS-IUI	3	292	Odds Ratio (M-H, Random, 95% CI)	2.61 [1.07, 6.37]

SUMMARY OF FINDINGS TABLES

1. Summary of findings - live birth or ongoing pregnancy

Estimates of effects, confidence intervals, and certainty of the evidence for live birth in couples with unexplained infertility

Patient or population: couples with unexplained infertility
Intervention: OS, IUI, OS-IUI, or IVF/ICSI
Comparator: expectant management, OS, IUI, or OS-IUI
Outcome: live birth
Setting: outpatient

All comparisons (10 RCTs, 2725 couples)		**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)****	**Quality of the evidence (GRADE)**
Comparator	**Intervention (number of RCTs and number of couples in direct comparison)**	**Assumed risk with comparator**	**Corresponding risk with intervention**		
Expectant management	OS (2 RCTs, 527 couples)	166 per 1000	**167 per 1000** (92 to 282)	**OR 1.01** (0.51 to 1.98)	⊕⊕⊖⊖ LOW[a]
	IUI (1 RCT, 386 couples)	166 per 1000	**194 per 1000** (108 to 325)	**OR 1.45** (0.61 to 2.43)	⊕⊕⊖⊖ LOW[a]
	OS-IUI (2 RCTs, 454 couples)	166 per 1000	**242 per 1000** (149 to 369)	**OR 1.61** (0.88 to 2.94)	⊕⊕⊖⊖ LOW[b]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	166 per 1000	**272 per 1000** (139 to 465)	**OR 1.88** (0.81 to 4.38)	⊕⊕⊖⊖ LOW[a]
OS	IUI (1 RCT, 387 couples)	174 per 1000	**201 per 1000** (107 to 346)	**OR 1.20** (0.57 to 2.52)	⊕⊕⊖⊖ LOW[a]
	OS-IUI (1 RCT, 184 couples)	174 per 1000	**252 per 1000** (145 to 399)	**OR 1.60** (0.81 to 3.16)	⊕⊕⊖⊖ LOW[a]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	174 per 1000	**281 per 1000** (136 to 492)	**OR 2.63** (0.75 to 4.61)	⊕⊕⊖⊖ LOW[a]
IUI	OS-IUI (2 RCTs, 636 couples)	166 per 1000	**209 per 1000** (128 to 323)	**OR 1.33** (0.67 to 3.58)	⊕⊕⊖⊖ LOW[a]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	166 per 1000	**235 per 1000** (117 to 416)	**OR 1.55** (0.67 to 3.58)	⊕⊕⊖⊖ LOW[a]

1. Summary of findings - live birth or ongoing pregnancy Continued

OS-IUI	IVF/ICSI (3 RCTs, 731 couples)	319 per 1000	**354 per 1000** (230 to 498)	**OR 1.17** (0.64 to 2.12)	⊕⊕⊖⊖ LOW[a]

CI: confidence interval; OR: odds ratio; RCT: randomised controlled trial.
***The corresponding risk in the intervention group** (and its 95% CI) is based on the mean risk in the comparator group and the relative effect of the intervention (and its 95% CI).
**All ORs and 95% CIs are based on network estimates.

GRADE Working Group grades of evidence.
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

Footnotes
[a]Downgraded by two levels for very serious imprecision.
[b]Downgraded by two levels for serious imprecision and serious heterogeneity.

2. Summary of findings - multiple pregnancy

Estimates of effects, confidence intervals, and certainty of the evidence for multiple pregnancy in couples with unexplained infertility

Patient or population: couples with unexplained infertility
Intervention: OS, OS-IUI, or IVF/ICSI
Comparator: expectant management/IUI, OS, or OS-IUI
Outcome: multiple pregnancy
Setting: outpatient

All comparisons (11 RCTs, 2564 couples)		**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)****	**Quality of the evidence (GRADE)**
Comparator	**Intervention (number of RCTs and number of couples in direct comparison)**	**Assumed risk with comparator**	**Corresponding risk with intervention**		
Expectant management/IUI	OS (3 RCTs, 934 couples)	6 per 1000	**17 per 1000** (6 to 50)	**OR 3.07** (1.00 to 9.41)	⊕⊕⊖⊖ LOW[a]
	OS-IUI (3 RCTs, 625 couples)	6 per 1000	**18 per 1000** (6 to 54)	**OR 3.34** (1.09 to 10.29)	⊕⊕⊕⊖ MODERATE[b]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	6 per 1000	**15 per 1000** (4 to 55)	**OR 2.66** (0.68 to 10.43)	⊕⊕⊖⊖ LOW[c]

2. Summary of findings - multiple pregnancy Continued

OS	OS-IUI (2 RCTs, 274 couples)	23 per 1000	**26 per 1000** (9 to 70)	**OR 1.09** (0.38 to 3.15)	⊕⊖⊖⊖ VERY LOW[d]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	23 per 1000	**20 per 1000** (6 to 72)	**OR 0.87** (0.23 to 3.24)	⊕⊕⊖⊖ LOW[c]
OS-IUI	IVF/ICSI (3 RCTs, 731 couples)	27 per 1000	**22 per 1000** (10 to 47)	**OR 0.80** (0.37 to 1.73)	⊕⊕⊖⊖ LOW[c]

CI: confidence interval; OR: odds ratio; RCT: randomised controlled trial.
***The corresponding risk in the intervention group** (and its 95% CI) is based on the mean risk in the comparator group and the relative effect of the intervention (and its 95% CI).
**All ORs and 95% CIs are based on network estimates.

GRADE Working Group grades of evidence.
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

Footnotes
[a]Downgraded by two levels for serious imprecision and serious heterogeneity.
[b]Downgraded by one level for serious imprecision.
[c]Downgraded by two levels for very serious imprecision.
[d]Downgraded by three levels for serious study limitations and very serious imprecision.

3. Summary of findings - clinical pregnancy

Estimates of effects, confidence intervals, and certainty of the evidence for clinical pregnancy in couples with unexplained infertility

Patient or population: couples with unexplained infertility
Intervention: OS, IUI, OS-IUI, or IVF/ICSI
Comparator: expectant management, OS, IUI, or OS-IUI
Outcome: clinical pregnancy
Setting: outpatient

All comparisons (23 RCTs, 3792 couples)		**Illustrative comparative risks* (95% CI)**		**Relative effect (95% CI)****	**Quality of the evidence (GRADE)**
Comparator	**Intervention (number of RCTs and number of couples in direct comparison)**	**Assumed risk with comparator**	**Corresponding risk with intervention**		
Expectant management	OS (6 RCTs, 939 couples)	157 per 1000	**234 per 1000** (155 to 337)	**OR 1.64** (0.99 to 2.73)	⊕⊖⊖⊖ VERY LOW[a]
	IUI (3 RCTs, 528 couples)	157 per 1000	**182 per 1000** (102 to 305)	**OR 1.20** (0.61 to 2.36)	⊕⊕⊖⊖ LOW[b]
	OS-IUI (4 RCTs, 525 couples)	157 per 1000	**301 per 1000** (205 to 420)	**OR 2.32** (1.39 to 3.90)	⊕⊕⊖⊖ LOW[c]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	157 per 1000	**360 per 1000** (197 to 563)	**OR 3.03** (1.32 to 6.94)	⊕⊕⊖⊖ LOW[c]
OS	IUI (2 RCTs, 407 couples)	213 per 1000	**165 per 1000** (93 to 277)	**OR 0.73** (0.38 to 1.42)	⊕⊖⊖⊖ VERY LOW[d]
	OS-IUI (8 RCTs, 763 couples)	213 per 1000	**276 per 1000** (199 to 371)	**OR 1.41** (0.92 to 2.18)	⊕⊖⊖⊖ VERY LOW[e]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	213 per 1000	**332 per 1000** (275 to 521)	**OR 1.84** (1.40 to 4.02)	⊕⊕⊖⊖ LOW[f]
IUI	OS-IUI (4 RCTs, 579 couples)	174 per 1000	**291 per 1000** (182 to 430)	**OR 1.94** (1.05 to 3.57)	⊕⊖⊖⊖ VERY LOW[a]
	IVF/ICSI (no direct evidence available; only indirect evidence used here)	174 per 1000	**347 per 1000** (180 to 566)	**OR 2.52** (1.04 to 6.16)	⊕⊕⊖⊖ LOW[f]

3. Summary of findings - clinical pregnancy Continued

OS-IUI	IVF/ICSI (3 RCTs, 731 couples)	344 per 1000	**437 per 1000** (289 to 599)	**OR 1.30** (0.68 to 2.50)	⊕⊕⊖⊖ LOW[b]

CI: confidence interval; OR: odds ratio; RCT: randomised controlled trial.
***The corresponding risk in the intervention group** (and its 95% CI) is based on the mean risk in the comparator group and the relative effect of the intervention (and its 95% CI).
**All ORs and 95% CIs are based on network estimates.

GRADE Working Group grades of evidence.
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

Footnotes
[a]Downgraded by three levels for serious study limitations, imprecision, and heterogeneity.
[b]Downgraded by two levels for very serious imprecision.
[c]Downgraded by two levels for very serious heterogeneity.
[d]Downgraded by three levels for very serious imprecision and serious incoherence.
[e]Downgraded by three levels for very serious study limitations, serious imprecision, and serious heterogeneity.
[f]Downgraded by two levels for serious imprecision and serious heterogeneity.

4. SUMMARY OF FINDINGS - MODERATE/SEVERE OHSS

IVF/ICSI compared with OS-IUI for unexplained infertility

Patient or population: couples with unexplained infertility
Settings: outpatient
Intervention: IVF/ICSI
Comparison: OS-IUI

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No. of participants (studies)	Quality of the evidence (GRADE)	Comments
	Assumed risk with OS-IUI	Corresponding risk with IVF/ ICSI				
Moderate/ severe OHSS	**11 per 1000**	**28 per 1000** (10 to 72)	**OR 2.50** (0.92 to 6.76)	958 (5 studies)	⊕⊕⊕⊖ MODERATE[a]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; OR: odds ratio; RCT: randomised controlled trial.

GRADE Working Group grades of evidence.
High quality: further research is very unlikely to change our confidence in the estimate of effect.
Moderate quality: further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
Low quality: further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
Very low quality: we are very uncertain about the estimate.

Footnotes
[a]Downgraded by one level for serious imprecision.

DISCUSSION

Summary of main results

This systematic review and network meta-analysis compared the effectiveness and safety of in vitro fertilisation (IVF)/intracytoplasmic sperm injection (ICSI), ovarian stimulation (OS)-intrauterine insemination (IUI), IUI, OS, and expectant management with each other in couples with unexplained infertility. There was insufficient evidence of differences in terms of live birth between expectant management and the other four interventions. Compared to expectant management or IUI, OS may increase the odds of multiple pregnancy, and OS-IUI probably increases the odds of multiple pregnancy. Evidence of differences between IVF/ICSI and expectant management for multiple pregnancy was insufficient. There was also insufficient evidence of a difference in moderate or severe ovarian hyperstimulation syndrome (OHSS) between IVF/ICSI and OS-IUI. The overall certainty of the evidence was low to moderate, mainly due to imprecision and/or heterogeneity.

Overall completeness and applicability of evidence

Our population of interest consisted of couples with unexplained infertility. We used a relatively broad definition of unexplained infertility, including couples with mild endometriosis and mild male infertility (pre-wash total motile sperm count > 3 * 10^6) to increase the applicability of findings. As the distributions of potential effect modifiers showed similarities across different comparisons and the interventions of interest are jointly randomisable, the overall transitivity assumption in this network was valid. For IVF/ICSI, all RCTs including this arm applied single embryo transfer policy, which guarantees the clinical homogeneity of IVF/ICSI.

Current guidelines (National Institute for Health and Care Excellence, 2013) do not recommend IUI, either with or without ovarian stimulation, for couples with unexplained infertility. Based on our systematic review, we would argue that OS-IUI still plays an important role in the treatment of unexplained infertility, especially for couples with poor prognosis of natural conception. Shared decision-making should consider not only effectiveness and safety, but also patient preferences and costs. Two economic evaluations found that OS-IUI resulted in lower cost per live birth than IVF/ICSI in couples with poor prognosis of natural conception and a median duration of infertility less than two years, which implies that OS-IUI is an important alternative to IVF/ICSI in these narrowly defined couples with unexplained infertility (Tjon-Kon-Fat, et al., 2015, van Rumste, et al., 2014).

Quality of the evidence

Overall certainty of the evidence was very low to moderate (Summary of findings table 1; Summary of findings table 2; Summary of findings table 3; Summary of findings table 4). This was due mainly to lack of precision and/or the existence of heterogeneity. All comparisons had relatively few included studies with direct evidence, which explained the imprecision in these comparisons. The heterogeneity observed was most likely

due to the heterogeneous nature of unexplained infertility, and some included RCTs focused on different subpopulations with unexplained infertility. For instance, (Steures, et al., 2006) included only couples with an intermediate prognosis of natural conception based on the Hunault prediction model (Hunault, et al., 2004), and (Farquhar, et al., 2017) included only couples with a poor prognosis. The result of network meta-analysis in the comparison of OS-IUI and expectant management was consistent with existing cohorts on unselected unexplained infertility (van Eekelen, et al., 2019), but the pooled result was not applicable to the two subpopulations with poor or intermediate prognoses, respectively.

The strengths of this systematic review include the extensive search strategy, use of indirect evidence, performance of sensitivity analyses, and application of Confidence in Network Meta-analysis (CINeMA) to evaluate the overall certainty of evidence in network meta-analysis. The current systematic review and network meta-analysis provided an overview of the evidence base in clinical management of unexplained infertility. Nevertheless, there are several limitations. Couples with unexplained infertility are a heterogeneous population, and various inclusion criteria were used. For instance, participants in the included studies may or may not have had a diagnostic laparoscopy before diagnosis of unexplained infertility. Next, some included studies focused on a subgroup of couples based on prognostic factors (e.g. Hunault prediction model as discussed above). Pooled results led to heterogeneity and imprecision in the evidence for these comparisons. Additionally, our primary effectiveness and safety outcomes live birth and multiple pregnancy were not reported in approximately half of the included trials. This explains in part the imprecision evident in some comparisons. Furthermore, as breakdown data for different subgroups were not available, our subgroup analysis on duration of infertility was based on different mean/median values; therefore these results should be interpreted with caution. A planned subgroup analysis on treatment-naive couples versus couples who had received prior treatment was not feasible in the network meta-analysis, as couples with various previous treatments were also allowed to be randomised to less invasive interventions, including expectant management in pragmatic RCTs. Last, about half of the included studies were published before 2000. Although IVF in different studies in this network meta-analysis appears similar, the intensive OS protocols and the relatively loose cancellation criteria used in old trials of OS and OS-IUI are not the same compared to recent ones, the latter of which led to fewer multiple pregnancies.

Potential biases in the review process

Given the extensive search strategy, including the electronic database search and the handsearch of relevant references, the chance of incomplete identification of studies was low. We did not identify small study effects in the main outcomes. Therefore, we concluded that no publication bias was evident. In addition, as live birth and/or multiple pregnancy was not reported in about half of the included studies, we could not rule out the possibility of reporting bias.

As indirect evidence does not involve new randomisation and therefore the validity of network meta-analysis relies on transitivity assumption, we assessed the transitivity assumption carefully before conducting this network meta-analysis and did not find evidence of intransitivity. However, we could not completely rule out the existence of intransitivity due to the small number of RCTs included in all comparisons and the lack of baseline information from old RCTs. We further evaluated inconsistency by using both global and local approaches. Statistical testing did not show evidence of inconsistency in networks of the main outcomes, but statistical testing for inconsistency could be underpowered (Higgins, et al., 2012). The overall limitations in each comparison on different outcomes are reflected in the summary of finding tables.

Agreements and disagreements with other studies or reviews

A Cochrane Review on IUI for unexplained infertility found no conclusive evidence of a difference in live birth or multiple pregnancy for the comparison between IUI or OS-IUI versus expectant management (Veltman-Verhulst, et al., 2016). Our network meta-analysis showed consistent results on live birth with overlapping confidence intervals. Evidence on multiple pregnancy between OS-IUI versus expectant management or IUI in our network meta-analysis was based on moderate certainty, as the use of network meta-analysis increased the precision of the evidence.

Another Cochrane Review on IVF/ICSI for unexplained infertility found that IVF/ICSI may be associated with higher live birth rates than expectant management, but the overall certainty of evidence was very low (Pandian, et al., 2015). This conclusion was based on one RCT with small sample size and an intensive embryo transfer policy (up to four embryos in an unselected population) (Hughes, et al., 2004). This RCT was not included in the network meta-analysis due to the different embryo transfer policy used from current clinical practice. No direct evidence was available for the comparison between IVF/ICSI and expectant management. Indirect evidence arising from our network meta-analysis was insufficient to judge a difference in terms of effectiveness and safety.

AUTHORS' CONCLUSIONS

Implications for practice

We found insufficient evidence of differences in terms of live birth between expectant management and the other four interventions (OS, IUI, OS-IUI, and IVF/ICSI). Compared to expectant management/IUI, OS may increase the odds of multiple pregnancy, and OS-IUI probably increases the odds of multiple pregnancy. Evidence showing differences between IVF/ICSI and expectant management for multiple pregnancy was insufficient, as was evidence of a difference in moderate or severe OHSS between IVF/ICSI and OS-IUI.

Implications for research

Given the overall low certainty of evidence for most comparisons in this network meta-analysis, future RCTs comparing interventions for unexplained infertility are needed. A recent systematic review showed that existing RCTs in reproductive medicine are likely to be underpowered to detect plausible improvements in live birth rate (Stocking, et al., 2019), as clinically important differences between these interventions appear small. Therefore, accounting for prognostic factors is helpful in guiding the design in future research. As the prognosis of natural conception in unexplained infertility is predicable, the relative effects between expectant management and other interventions are expected to be larger in couples with poor prognosis. This was confirmed not only in our subgroup analysis, which showed different effects in couples with shorter and longer duration of infertility, but also in our sensitivity analysis, which showed large relative effects in couples with poor prognosis. Future RCTs should compare IVF or OS-IUI versus expectant management in couples with different prognoses to confirm the available evidence and to shape the clinical indications for IVF and IUI in unexplained infertility.

We need more studies comparing OS-IUI or IVF versus expectant management as well as studies comparing OS-IUI versus IVF to enable better fine-tuning of when to start treatment and what treatment to use. More specifically, in an OS-IUI protocol, gonadotropins with strict cancellation criteria and recently widely used medication such as letrozole should be tested. Studies comparing IVF versus other interventions should also address the use of the freeze-only strategy and the report of cumulative live birth rate.

Studies should include a cost-effectiveness analysis with a time horizon that allows multi-cycle treatment plus frozen-thawed cycles in cases of IVF, with live birth as the primary outcome.

Study investigators are advised to use cumulative live birth as the primary outcome. Cumulative live birth has been recognised as the current standard in outcome reporting (Gadalla, et al., 2018). The development of a core outcome set for infertility trials is under way (Duffy, et al., 2018). The use of core outcomes will standardise outcome reporting in future trials and will minimise outcome reporting bias.

ACKNOWLEDGEMENTS

We would like to thank Marian Showell from the Cochrane Gynaecology and Fertility Group for conducting the database searches.

Contributions of authors

RW, RIT, MJCE, PMMB, FvdV, SB, BWM, and MvW contributed to the study design and the original protocol. RW, NAD, RIT, and MvW collected the data. RW and MvW analysed the data. All review authors interpreted the data. RW wrote the first draft.

All review authors revised the manuscript critically for important intellectual content and approved the final version.

Declarations of interest

SB has not received money from any source to support the work leading up to this review. SB has received support for travel and accommodation for speaking at conferences. His institution and institutional colleagues have received support from pharmaceutical companies for educational activities such as hosting seminars and attending conferences. He receives an honorarium as Editor-in-Chief of *Human Reproduction Open*.

BWM is supported by an NHMRC Practitioner Fellowship (GNT1082548) and reports consultancy for ObsEva, Merck, and Guerbet, and research grants from Guerbet and Merck.

RW, NAD, RIT, ME, PB, MHM, FvdV, and MvW have no interests to declare.

Differences between protocol and review

We replaced subfertility with infertility according to the latest version of the International Glossary on Infertility and Fertility Care (Zegers-Hochschild 2017). We excluded studies on modified natural cycle IVF as it is different from IVF with ovarian hyperstimulation.

We planned in the protocol to perform a sensitivity analysis by using alternative imputation strategies. However, for binary outcomes, it can be problematic to impute missing outcomes as events. Therefore, we did a sensitivity analysis by excluding missing outcome data as recommended by the *Cochrane Handbook for Systematic Reviews of Interventions*. We did not report the predictive interval in this network meta-analysis but used it when accessing heterogeneity for the overall certainty of evidence in CINeMA (CINeMA 2017; Salanti 2014).

SUPPLEMENTARY DATA

Supplementary data are available online

REFERENCES

Agarwal S, Mittal S. A randomised prospective trial of intrauterine insemination versus timed intercourse in superovulated cycles with clomiphene. *Indian Journal of Medical Research* 2004;120: 519-522.

American Society for Reproductive Medicine. Effectiveness and Treatment for Unexplained Infertility *The practice committee of the American Society for Reproductive Medicine*. 2006.

Arcaini L, Bianchi S, Baglioni A, Marchini M, Tozzi L, Fedele L. Superovulation and intrauterine insemination vs. superovulation alone in the treatment of unexplained infertility. A randomized study. *Journal of Reproductive Medicine* 1996;41: 614-618.

Arici A, Byrd W, Bradshaw K, Kutteh WH, Marshburn P, Carr BR. Evaluation of clomiphene citrate and human chorionic gonadotropin treatment: a prospective, randomized, crossover study during intrauterine insemination cycles. *Fertility & Sterility* 1994;61: 314-318.

Athaullah N, Proctor M, Johnson N. Oral versus injectable ovulation induction agents for unexplained subfertility. *Cochrane Database of Systematic Reviews* 2002.

Bensdorp AJ, Tjon-Kon-Fat RI, Bossuyt PMM, Koks CAM, Oosterhuis GJE, et al. Prevention of multiple pregnancies in couples with unexplained or mild male subfertility: randomised controlled trial of in vitro fertilisation with single embryo transfer or in vitro fertilisation in modified natural cycle compared with intrauterine insemination with controlled ovarian hyperstimulation. *BMJ (Online)* 2015;350: g7771-g7771.

Bhattacharya S, Harrild K, Mollison J, Wordsworth S, Tay C, Harrold A, et al. Clomifene citrate or unstimulated intrauterine insemination compared with expectant management for unexplained infertility: pragmatic randomised controlled trial. *BMJ* 2008;337: a716-a716.

Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Human Reproduction* 2007;22: 1506-1512.

Brandes M, Hamilton CJCM, de Bruin JP, Nelen WLDM, Kremer JAM. The relative contribution of IVF to the total ongoing pregnancy rate in a subfertile cohort. *Human Reproduction* 2010;25: 118-126.

Brandes M, Hamilton CJCM, van der Steen JOM, de Bruin JP, Bots RSGM, Nelen WLDM, et al. Unexplained infertility: overall ongoing pregnancy rate and mode of conception. *Human Reproduction* 2011;26: 360-368.

Buvat J, Buvat Herbaut M, Marcolin G, Guittard C, Herbaut M, Verbecq P. Ovarian hyperstimulation and insemination vs IVF in unexplained infertility: results of randomized trial (poster). *Contraception, Fertilite, Sexualite* 1993;21: 432-432.

Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS ONE* 2013;8: e76654-e76654.

Chaimani A, Salanti G. Visualizing assumptions and results in network meta-analysis: the network graphs package. *Stata Journal* 2015;15: 905-950.

Chung CC, Fleming R, Jamieson ME, Yates RW, Coutts JR. Randomized comparison of ovulation induction with and without intrauterine insemination in the treatment of unexplained infertility. *Human Reproduction* 1995;10: 3139-3141.

CINeMA. CINeMA: Confidence in Network Meta-Analysis [Software]. 2017. Institute of Social and Preventive Medicine, University of Bern 2017. Available from cinema.ispm.ch.

Crosignani P, Walters DE, Soliani A. The ESHRE multicentre trial on the treatment of unexplained infertility. *Fertility & Sterility* 1991;54: S61-S61.

Crosignani PG, Walters DE, Soliani A. The ESHRE multicentre trial on the treatment of unexplained infertility: a preliminary report. European Society of Human Reproduction and Embryology. *Human Reproduction* 1991;6: 953-958.

Custers IM, Konig TE, Broekmans FJ, Hompes PG, Kaaijk E, Oosterhuis J, et al. Couples with unexplained subfertility and unfavorable prognosis: a randomized pilot trial comparing the effectiveness of in vitro fertilization with elective single embryo transfer versus intrauterine insemination with controlled ovarian stimulation. *Fertility & Sterility* 2011;96: 1107-1111. e1101.

Datta J, Palmer MJ, Tanton C, Gibson LJ, Jones KG, Macdowall W, et al. Prevalence of infertility and help seeking among 15 000 women and men. *Human Reproduction* 2016;31: 2108-2118.

Deaton JL, Gibson M, Blackmer KM, Nakajima ST, Badger GJ, Brumsted JR. A randomized, controlled trial of clomiphene citrate and intrauterine insemination in couples with unexplained infertility or surgically corrected endometriosis. *Fertility & Sterility* 1990;54: 1083-1088.

Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;29: 932-944.

Duffy JMN, Bhattacharya S, Curtis C, Evers JLH, Farquharson RG, et al. A protocol developing, disseminating and implementing a core outcome set for infertility. *Human Reproduction Open* 2018;2018: hoy007-hoy007.

Dutch Society of Obstetrics and Gynaecology. Guideline: unexplained subfertility *Dutch Society of Obstetrics and Gynaecology (NVOG)*. 2010.

Elzeiny H, Garrett C, Toledo M, Stern K, McBain J, Baker HW. A randomised controlled trial of intra-uterine insemination versus in vitro fertilisation in patients with idiopathic or mild male infertility. *Australian & New Zealand Journal of Obstetrics & Gynaecology* 2014;54: 156-161.

Farquhar CM, Liu E, Armstrong S, Arroll N, Lensen S, Brown J. Intrauterine insemination with ovarian stimulation versus expectant management for unexplained infertility (TUI): a pragmatic, open-label, randomised, controlled, two-centre trial. *Lancet* 2017;391: 441-450.

Fisch P, Casper RF, Brown SE, Wrixon W, Collins JA, Reid RL, et al. Unexplained infertility: evaluation of treatment with clomiphene citrate and human chorionic gonadotropin. *Fertility & Sterility* 1989;51: 828-833.

Fujii S, Fukui A, Fukushi Y, Kagiya A, Sato S, Saito Y. The effects of clomiphene citrate on normally ovulatory women. *Fertility & Sterility* 1997;68: 997-999.

Gadalla MA, Wang R, van Wely M, Mol BWJ. How should we report outcomes in reproductive medicine? *Ultrasound in Obstetrics and Gynecology* 2018;51: 7-9.

George K, George SS, Chandy A. A randomized controlled trial to evaluate the efficacy of clomiphene citrate in unexplained infertility. *Fertility and Sterility* 2006;86 Suppl 2: S100-S100.

Glazener CM, Coulson C, Lambert PA, Watt EM, Hinton RA, Kelly NG, et al. Clomiphene treatment for women with unexplained infertility: placebo-controlled study of hormonal responses and conception rates. *Gynecological Endocrinology* 1990;4: 75-83.

Gnoth C. Time to pregnancy: results of the German prospective study and impact of the management of infertility. *Human Reproduction* 2003;18: 1959-1966.

Goldman MB, Regan MM, Alper MM, Thornton KL, Reindollar RH. Pregnancy rates across multiple treatment cycles: data from the fast track and standard treatment (FASTT) trial. *Fertility and Sterility* 2010;94 Suppl 1: S1 Abstract no. O-S1 Abstract no.01.

Goldman MB, Thornton KL, Ryley D, Alper MM, Fung JL, Hornstein MD, et al. A randomized clinical trial to determine optimal infertility treatment in older couples: the Forty and Over Treatment Trial (FORT-T). *Fertility & Sterility* 2014;101: 1574-1581.e1571-1572.

Goverde AJ, McDonnell J, Vermeiden JP, Schats R, Rutten FF, Schoemaker J. Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis. *Lancet* 2000;355: 13-18.

Gregoriou O, Vitoratos N, Papadias C, Konidaris S, Gargaropoulos A, Louridas C. Controlled ovarian hyperstimulation with or without intrauterine insemination for the treatment of unexplained infertility. *International Journal of Gynaecology & Obstetrics* 1995;48: 55-59.

Gunn DD, Bates GW. Evidence-based approach to unexplained infertility: a systematic review. *Fertility and Sterility* 2016;105: 1566-1574.

Guzick DS, Carson SA, Coutifaris C, Overstreet JW, Factor-Litvak P, Steinkampf MP, et al. Efficacy of superovulation and intrauterine insemination in the treatment of infertility. National Cooperative Reproductive Medicine Network. *New England Journal of Medicine* 1999;340: 177-183.

Harrison R, O'Moore R. The use of clomiphene citrate with and without human chorionic gonadotropin. *Irish Medical Journal* 1983;76: 273-274.

Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods* 2012;3: 98-110.

Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011 *Available from handbookcochraneorg*. 2011.

Ho PC, Yeung WSB, So WWK, Lau EYL. A randomised trial comparing the efficacy of ovarian stimulation and intrauterine insemination versus ovarian stimulation alone in the treatment of male infertility and unexplained infertility. *British Journal of Obstetrics & Gynaecology* 1998;105: 43-43.

Huang S, Wang R, Li R, Wang H, Qiao J, Mol BWJ. Ovarian stimulation in infertile women treated with the use of intrauterine insemination: a cohort study from China. *Fertility and Sterility* 2018;109: 872-878.

Hughes E, Brown J, Collins JJ, Vanderkerchove P. Clomiphene citrate for unexplained subfertility in women. *Cochrane Database of Systematic Reviews* 2010.

Hughes EG, Beecroft ML, Wilkie V, Burville L, Claman P, Tummon I, et al. A multicentre randomized controlled trial of expectant management versus IVF in women with Fallopian tube patency. *Human Reproduction (Oxford, England)* 2004;19: 1105-1109.

Hull MG, Glazener CM, Kelly NJ, Conway DI, Foster PA, Hinton RA, et al. Population study of causes, treatment, and outcome of infertility. *British Medical Journal (Clinical Research Ed)* 1985;291: 1693-1697.

Hunault CC, Habbema JD, Eijkemans MJ, Collins JA, Evers JL, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. *Human Reproduction* 2004;19: 2019-2026.

Hunault CC, Laven JS, van Rooij IA, Eijkemans MJ, te Velde ER, Habbema JD. Prospective validation of two models predicting pregnancy leading to live birth among untreated subfertile couples. *Human Reproduction* 2005;20: 1636-1641.

Janko P, Hruzik P, Pruzinec J, Saliba H, Zidzik J. Induction of ovulation with or without intrauterine insemination in cases of unexplained sterility. *Fertility and Sterility* 1998;70: S442-S442.

Kamphuis EI, Bhattacharya S, van der Veen F, Mol BW, Templeton A, Evidence Based IVFG. Are we overusing IVF? *BMJ* 2014;348: g252-g252.

Kandavel V, Cheong Y. Does intra-uterine insemination have a place in modern ART practice? *Best Practice & Research: Clinical Obstetrics & Gynaecology* 2018;53: 3-10.

Karlstrom PO, Bergh T, Lundkvist O. A prospective randomized trial of artificial insemination versus intercourse in cycles stimulated with human menopausal gonadotropin or clomiphene citrate. *Fertility & Sterility* 1993;59: 554-559.

Kirby CA, Flaherty SP, Godfrey BM, Warnes GM, Matthews CD. A prospective trial of intrauterine insemination of motile spermatozoa versus timed intercourse. *Fertility & Sterility* 1991;56: 102-107.

Leanza V, Coco L, Grasso F, Leanza G, Zarbo G, Palumbo M. Ovulation induction with clomiphene citrate for infertile couple. *Minerva Ginecologica* 2014;66: 309-312.

Leanza V, Coco L, Grasso F, Leanza G, Zarbo G, Palumbo M. Unexplained infertility and ovulatory induction with menopausal gonadotropins. *Minerva Ginecologica* 2014;66: 303-307.

Martinez AR, Bernardus RE, Voorhorst FJ, Vermeiden JP, Schoemaker J. Intrauterine insemination does and clomiphene citrate does not improve fecundity in couples with infertility due to male or idiopathic factors: a prospective, randomized, controlled study. *Fertility & Sterility* 1990;53: 847-853.

Martinez AR, Bernardus RE, Voorhorst FJ, Vermeiden JP, Schoemaker J. Pregnancy rates after timed intercourse or intrauterine insemination after human menopausal gonadotropin stimulation of normal ovulatory cycles: a controlled study. *Fertility & Sterility* 1991;55: 258-265.

Mathur R, Evbuomwan I, Jenkins J. Prevention and management of ovarian hyperstimulation syndrome. *Current Obstetrics and Gynaecology* 2005;15: 132-138.

Melis G, Paoletti A, Strigini F, Fabris F, Canale D, Fioretti P. Pharmacologic induction of multiple follicular development improves the success rate of artificial insemination with husband's semen in couples with male-related or unexplained infertility. *Fertility & Sterility* 1987;47: 441-445.

Melis GB, Paoletti AM, Ajossa S, Guerriero S, Depau GF, Mais V. Ovulation induction with gonadotropins as sole treatment in infertile couples with open tubes: a randomized prospective comparison between intrauterine insemination and timed vaginal intercourse. *Fertility & Sterility* 1995;64: 1088-1093.

Murdoch AP, Harris M, Mahroo M, Williams M, Dunlop W. Gamete intrafallopian transfer (GIFT) compared with intrauterine insemination in the treatment of unexplained infertility. *British Journal of Obstetrics & Gynaecology* 1991;98: 1107-1111.

Nandi A, Bhide P, Hooper R, Gudi A, Shah A, Khan K, et al. Intrauterine insemination with gonadotropin stimulation or in vitro fertilization for the treatment of unexplained subfertility: a randomized controlled trial. *Fertility & Sterility* 2017;107: 1329-1335.e1322.

National Institute for Health and Care Excellence. Fertility problems: assessment and treatment, Clinical guideline [CG156]. 2013.

NCT01992731. IUI vs. IVF/ICSI in Women Aged 38-42 Years: A Prospective Randomized Controlled Trial *clinicaltrialsgov/ct2/show/NCT01992731*. 2013.

NCT02001870. Comparison of the Efficiency of Intra-uterine Insemination and in Vitro Fertilization in Women Over 37 Years (AMPAGE) *clinicaltrialsgov/ct2/show/NCT02001870*. 2013.

NCT02461173. Stimulated Intrauterine Insemination Cycles and Unstimulated Intrauterine Insemination Cycles in Couples With Unexplained Infertility *clinicaltrialsgov/ct2/show/NCT02461173*. 2015.

NCT03455426. Intrauterine Insemination With Letrozole Versus Intrauterine Insemination in Natural Cycle. A Randomised Controlled Trial *clinicaltrialsgov/ct2/show/NCT03455426*. 2018.

NTR5599. Intrauterine Insemination for Unexplained or Mild Male Subfertility *trialregisternl/trialreg/admin/rctviewasp?TC=5599*. 2016.

Nulsen JC, Walsh S, Dumez S, Metzger DA. A randomized and longitudinal study of human menopausal gonadotropin with intrauterine insemination in the treatment of infertility. *Obstetrics & Gynecology* 1993;82: 780-786.

Palermo G, Joris H, Devroey P, Van Steirteghem AC. Pregnancies after intracytoplasmic injection of single spermatozoon into an oocyte. *Lancet* 1992;340: 17-18.

Pandian Z, Gibreel A, Bhattacharya S. In vitro fertilisation for unexplained subfertility. *Cochrane Database of Systematic Reviews* 2015.

Prentice A, Sacks GP, Morton NC, Deary AJ, Smith SK. Controlled ovarian stimulation (superovulation) and intrauterine insemination for the treatment of unexplained and minor male factor infertility. *Human Reproduction (Oxford, England)* 1995;10: 112-112.

Reindollar RH, Regan MM, Neumann PJ, Levine BS, Thornton KL, Alper MM, et al. A randomized clinical trial to evaluate optimal treatment for unexplained infertility: the fast track and standard treatment (FASTT) trial. *Fertility & Sterility* 2010;94: 888-899.

Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology* 2011;64: 163–171-163–171.

Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014;9: e99682-e99682.

Shokeir TA. Tamoxifen citrate for women with unexplained infertility. *Archives of Gynecology & Obstetrics* 2006;274: 279-283.

Soliman S, Daya S, Collins J, Jarrell J. A randomized trial of in vitro fertilization versus conventional treatment for infertility. *Fertility and Sterility* 1993;59: 1239-1244.

Steptoe PC, Edwards RG. Birth after the reimplantation of a human embryo. *Lancet* 1978;2: 366-366.

Steures P, van der Steeg JW, Hompes PG, Habbema JD, Eijkemans MJ, Broekmans FJ, Collaborative Effort on the Clinical Evaluation in Reproductive Medicine. Intrauterine insemination with controlled ovarian hyperstimulation versus expectant management for couples with unexplained subfertility and an intermediate prognosis: a randomised clinical trial. *Lancet* 2006;368: 216-221.

Stocking K, Wilkinson J, Lensen S, Brison DR, Roberts SA, Vail A. Are interventions in reproductive medicine assessed for plausible and clinically relevant effects? A systematic review of power and precision in trials and meta-analyses. *Human Reproduction* 2019;34: 659-665.

Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A, et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? *Human Reproduction* 2015;30: 2331-2339.

Tjon-Kon-Fat RI, Bensdorp AJ, Mol BW, van der Veen F, van Wely M. The natural conception rate in couples with unexplained or mild male subfertility scheduled for treatment with IVF-SET, IVF-MNC or IUI-COH (INeS trial). *Human Reproduction* 2014;29 suppl 1: i214 Abstract no: P-i214 Abstract no234.

van Eekelen R, Scholten I, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, van Wely M, et al. Natural conception: repeated predictions over time. *Human Reproduction* 2017;32: 346-353.

van Eekelen R, van Geloven N, van Wely M, McLernon DJ, Mol F, Custers IM, et al. Is IUI with ovarian stimulation effective in couples with unexplained subfertility? *Human Reproduction* 2019;34: 84-91.

van Rumste MM, Custers IM, van Wely M, Koks CA, van Weering HG, Beckers NG, et al. IVF with planned single-embryo transfer versus IUI with ovarian stimulation in couples with unexplained subfertility: an economic analysis. *Reproductive Biomedicine Online* 2014;28: 336-342.

Veltman-Verhulst SM, Hughes E, Olugbenga Ayeleke R, Cohlen BJ. Intra-uterine insemination for unexplained subfertility. *Cochrane Database of Systematic Reviews* 2016.

White IR. Network meta-analysis. *Stata Journal* 2015;15: 951-985.

Zayed F, Lenton EA, Cooke ID. Comparison between stimulated in-vitro fertilization and stimulated intrauterine insemination for the treatment of unexplained and mild male factor infertility. *Human Reproduction* 1997;12: 2408-2413.

Zegers-Hochschild F, Adamson GD, Dyer S, Racowsky C, de Mouzon J, Sokol R, et al. The International Glossary on Infertility and Fertility Care, 2017. *Human Reproduction* 2017;32: 1786-1801.

Zikopoulos K, West CP, Thong PW, Kacser EM, Morrison J, Wu FC. Homologous intra-uterine insemination has no advantage over timed natural intercourse when used in combination with ovulation induction for the treatment of unexplained infertility. *Human Reproduction* 1993;8: 563-567.

Zolghadri J, Younesi M, Tabibi A, Khosravi D, Behdin S, Vafayee H. Comparison of intrauterine insemination with timed intercourse method in controlled ovarian hyperstimulation. *Iranian Journal of Reproductive Medicine* 2012;10: 42-42.

3.

Follicle stimulating hormone (FSH) versus Clomiphene Citrate (CC) in intrauterine insemination for unexplained subfertility: A randomized controlled trial

authors:

N.A. Danhof, M. van Wely, S. Repping, C. Koks, H.R. Verhoeve, J.P. de Bruin, M.F.G. Verberg, M.H.A. van Hooff, B.J. Cohlen, C.F. van Heteren, K. Fleischer, J. Gianotten1, J. van Disseldorp, J. Visser, F.J.M. Broekmans, B.W.J. Mol, F. van der Veen, M.H. Mochtar, for the SUPER study group

Human Reproduction 2018;33(10):1866-74

ABSTRACT

Study question: Is follicle stimulation hormone (FSH) or clomiphene citrate (CC) the most effective stimulation regimen in terms of ongoing pregnancies in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria as a measure to reduce the number of multiple pregnancies?

Summary answer: In IUI with adherence to strict cancellation criteria, ovarian stimulation with FSH is not superior to CC in terms of the cumulative ongoing pregnancy rate, and yields a similar, low multiple pregnancy rate.

What is already known: FSH has been shown to result in higher pregnancy rates compared to CC, but at the cost of high multiple pregnancy rates. To reduce the risk of multiple pregnancy, new ovarian stimulation regimens have been suggested, these include strict cancellation criteria to limit the number of dominant follicles per cycle i.e. withholding insemination when more than three dominant follicles develop. With such a strategy, it is unclear whether the ovarian stimulation should be done with FSH or with CC.

Study design, size, duration: We performed an open-label multicenter randomized superiority controlled trial in the Netherlands (NTR 4057).

Participants/materials, setting, methods: We randomized couples diagnosed with unexplained subfertility and scheduled for a maximum of four cycles of IUI with ovarian stimulation with 75 IU FSH or 100 mg CC. Cycles were cancelled when more than three dominant follicles developed. The primary outcome was cumulative ongoing pregnancy rate. Multiple pregnancy was a secondary outcome. We analysed the data on intention to treat basis. We calculated relative risks and absolute risk difference with 95% CI.

Main results and the role of chance: Between July 2013 and March 2016, we allocated 369 women to ovarian stimulation with FSH and 369 women to ovarian stimulation with CC. A total of 113 women (31%) had an ongoing pregnancy following ovarian stimulation with FSH and 97 women (26%) had an ongoing pregnancy following ovarian stimulation with CC (RR 1.16, 95% CI: 0.93 to 1.47, ARD = 0.04, 95% CI: −0.02 to 0.11). Five women (1.4%) had a multiple pregnancy following ovarian stimulation with FSH and eight women (2.2%) had a multiple pregnancy following ovarian stimulation with CC (RR 0.63, 95% CI: 0.21 to 1.89, ARD = −0.01, 95% CI: −0.03 to 0.01).

Limitations, reasons for caution: We were not able to blind this study due to the nature of the interventions. We consider it unlikely that this has introduced performance bias, since pregnancy outcomes are objective outcome measures.

Wider implications of the findings: We revealed that adherence to strict cancellation criteria is a successful solution to reduce the number of multiple pregnancies in IUI. To decide whether ovarian stimulation with FSH or with CC should be the regimen of choice, costs and patients' preferences should be taken into account.

Study funding/competing interests: This trial received funding from the Dutch Organization for Health Research and Development (ZonMw). Prof. Dr B.W.J. Mol is supported by a NHMRC Practitioner Fellowship (GNT1082548). B.W.M. reports consultancy for Merck, ObsEva and Guerbet. The other authors declare that they have no competing interests.

INTRODUCTION

Annually, more than 70 million couples worldwide fail to conceive within 1 year of regular unprotected intercourse (Boivin et al., 2007). At present, in many countries the first line treatment for couples diagnosed with unexplained subfertility is IUI with ovarian stimulation (Calhaz-Jorge et al., 2017; The Practice Committee of the American Society for Reproductive Medicine., 2006). The downside of ovarian stimulation is the high multiple pregnancy risk with its increased risk of serious neonatal morbidity, neonatal mortality and maternal morbidity (Guzick et al.,1999; Ombelet et al., 2006).

According to a Cochrane review published in 2007, follicle stimulating hormone (FSH) is the drug of choice (Cantineau and Cohlen, 2007). The meta-analysis showed statistically significant increased pregnancy rates in favour of FSH compared to ovarian stimulation with clomiphene citrate (CC) in women undergoing IUI (seven studies, 556 women, OR 1.8, 95% CI: 1.2 to 2.7), while the -limited- data on multiple pregnancy rates were similar between FSH and CC and not allowing any conclusions (three studies, 338 women, OR 0.53, 95% CI: 0.15 to 1.86) (Cantineau and Cohlen, 2007). Since then, a recent large RCT, comparing FSH with CC in IUI also showed a statistically significant increase in live birth rates compared to CC, but at the cost of 25 twins and six triplets among 301 women (10%) undergoing ovarian stimulation with FSH, while there were 8 twins among 300 women (3%) undergoing ovarian stimulation with CC (Diamond et al., 2015). These high multiple pregnancy rates are no longer acceptable in modern infertility treatment.

To reduce the risk on multiple pregnancy, new ovarian stimulation regimens have been suggested, the quintessence of which are strict cancellation criteria to limit the number of dominant follicles per cycle, i.e. withholding insemination when more than three dominant follicles develop (Rumste van et al., 2006; Rumste van et al., 2008). The Cochrane review included one study that compared FSH to CC in a stimulation regimen with adherence to strict cancellation criteria (Dankert et al., 2007). This study found similar pregnancy rates (34% for FSH versus 38% for CC, RR 0.90, 95% CI: 0.57 to 1.41) and low multiple pregnancy rates (4% per cycle for FSH versus 7% per cycle for CC, RR 0.63, 95% CI: 0.06 to 6.53), but with only 138 included couples this study was underpowered (Dankert et al., 2007).

We therefore aimed to study, in a well powered randomized clinical trial the effectiveness of ovarian stimulation with 75 IU FSH compared to ovarian stimulation with 100 mg CC, in an IUI programme with adherence to strict cancellation criteria, i.e. cancellation of the cycle when more than three dominant follicles develop in women undergoing IUI, within a time horizon of 6 months.

MATERIALS AND METHODS

Study design

This study was an open-label multicenter, randomized controlled superiority trial positioned in the Dutch Consortium for Healthcare Evaluation in Obstetrics and Gynaecology (https://zorgevaluatienederland.nl/associations/1). We recruited couples between July 2013 and March 2016. The Medical Ethical Committee of the Academic Medical Centre and the Dutch Central Committee on Research involving Human Subjects approved this study (CCMO NL 43131-018-13) and the board of directors of each participating site approved local execution (NTR4057). The protocol (see Supplementary material) was published previously (Danhof et al., 2017).

Study population

Couples diagnosed with unexplained subfertility were eligible for the study. Unexplained subfertility was defined as a failure to conceive after one year of regular unprotected intercourse and a prewash total motile sperm count (TMSC) of at least 3 million (NICE clinical guideline). The inclusion criteria were female age between 18 and 43 years, regular menstrual cycle, at least one side tube patency and a TMSC of at least 3 million (NICE clinical guideline). If women were under 38 years of age, their 12 months prognosis on natural conception according to the model of Hunault had to be lower than 30% (Hunault et al., 2004; Steeg van der et al., 2007). Women were also eligible for inclusion after 6 months of failed expectant management. Women undergoing donor sperm treatment were eligible if they were below 35 years of age, had a regular menstrual cycle, with a least one-sided tubal patency, and had had 12 months of failed intracervical or IUI without ovarian stimulation or were above 35 years of age, had a regular menstrual cycle, with a least one-sided tubal patency and had 6 months of failed intracervical or IUI without ovarian stimulation.

Women with double sided tubal pathology, polycystic ovary syndrome, irregular cycles or other endocrine disorders were not eligible.

Interventions

We treated couples for a maximum of four cycles or until pregnancy occurred within a time horizon of 6 months. In the first treatment cycle, all women were seen for a baseline visit for a transvaginal ultrasound examination on the third, fourth or fifth day of the menstrual cycle. Women were not allowed to start the treatment cycle if one or more ovarian cysts of >20 mm were seen. In the experimental arm women started with daily subcutaneous injections of 75 IU FSH on Day 3, 4 or 5 of the menstrual cycle and continued these injections until the day of ovulation triggering (Dankert et al., 2007). In the standard arm women started with 100 mg CC on Day 3, 4 or 5 of the menstrual cycle. The tablets were administered orally and stopped after 5 days of daily intake.

In both interventions, we monitored follicular development by transvaginal ultrasound. We triggered ovulation with 5000 IU hCG or with 250 µg recHCG if there was at least one dominant follicle with a mean diameter of 16–18 mm a maximum of three follicles of

≥14 mm. At the final ultra- sound examination before ovulation triggering, we measured the total number of follicles their diameters and the endometrial thickness. We cancelled the cycle if more than three follicles with a diameter of ≥14 mm or five follicles with a diameter of ≥12 mm was seen at transvaginal ultra- sound, regardless of the endometrial thickness. In these cycles, we advised the couples to have protected or no intercourse. We scheduled IUI 36–42 h after ovulation triggering. On the day of insemination, the partner provided a semen sample after a minimum of 2 days of sexual abstinence. The semen was processed according to local protocol. In case of donor sperm treatment, donor semen was thawed and processed according to local protocol.

Women who did not conceive were scheduled for the next insemination cycle. In case of monofollicular growth, the dose of FSH was increased by 37.5 IU per day or the dose of CC was increased by 50 mg per day in the next cycle. If a cycle was cancelled due to the development of more than three dominant follicles, the dose of FSH was decreased by 37.5 IU per day or the dose of CC was decreased by 50 mg per day in the next cycle.

We treated couples for a maximum of four cycles or until pregnancy occurred within a time horizon of 6 months.

Clinical and ongoing pregnancies were confirmed by ultrasound.

Outcome measures

The primary outcome was ongoing pregnancy per woman, defined as a positive heartbeat at or beyond 12 weeks of gestation. Pregnancies that occurred within the first 6 months after randomization counted for assessment of the primary outcome.

Secondary outcomes per started cycle were cancellation rates, number of cycles with a single follicle, total number of follicles ≥14 mm at the time of ovulation triggering, and secondary outcomes per women were multiple pregnancy defined as registered heartbeat of at least two fetuses at 12 weeks of gestation, time to ongoing pregnancy, clinical pregnancy, defined as any registered foetal heartbeat on ultrasound, miscarriage, defined as pregnancy loss at a gestational age of 20 weeks or less, ectopic pregnancy and live birth.

Serious adverse events were reported to the trial coordinator.

Sample size calculation

We designed the study as a superiority trial. In our original sample size, we assumed the ongoing pregnancy rate was 35% after a maximum of 4 months of IUI with ovarian stimulation with CC. To be able to show a difference of 17.5% between ovarian stimulation with FSH and CC, we needed to recruit 182 couples per treatment arm with a two-sided alpha of 5% and a beta of 20%. Accounting for a 10% drop-out rate, we needed to recruit 404 women. In May 2015, we extended the sample size based on new available data. We applied an ongoing pregnancy rate of 25% following CC after four cycles and within 6 months (Bensdorp et al., 2015). To be able to show a minimally clinical relevant difference of 10% between ovarian stimulation with FSH and CC, we

needed to recruit 329 couples per treatment arm with a two-sided alpha of 5% and a beta of 20%. Accounting for a 10% drop-out rate, we needed to recruit 732 women.

Randomization and masking

Eligible women were informed about the study by their doctor or by a dedicated research nurse. After written informed consent women were randomized using a central password protected Internet-based randomization programme. The randomization list had been prepared by an independent statistician with a variable block size with randomly selected block sizes that varied between two, four and six. There was no stratification. Neither the recruiters nor the trial project group could access the randomization sequence.

Statistical analysis

We analysed all outcomes on an intention to treat basis. We also performed a per protocol analysis for the primary outcome and time to ongoing pregnancy, which was not pre-planned. We expressed all out- comes per couple randomized unless otherwise stated. We estimated differences in the primary and secondary outcomes as relative risks and absolute risk difference with 95% CI and used a Chi square test for formal analysis. We assessed the association between multiple pregnancy and follicle count using logistic regression models. We constructed Kaplan–Meier curves for the time to ongoing pregnancy. Pregnancies were timed at conception and a few women had undetected spontaneous pregnancies at randomization. These were included in the intention to treat analysis and appear as pregnancies at zero time (Lachin, 2000). We considered P values below 0.05 to indicate statistical significance.

Study oversight and role of the funding source

This trial was funded by the Netherlands Organization for Health Research and Development (ZonMw) (Health Care Efficiency Research; project number 80-83600-98-10192). The sponsor of the study had no role in study design, data collection, data analysis, data interpretation or writing the report. The corresponding author confirms to have had full access to all the data in the study and had final responsibility for the decision to submit for publication.

RESULTS

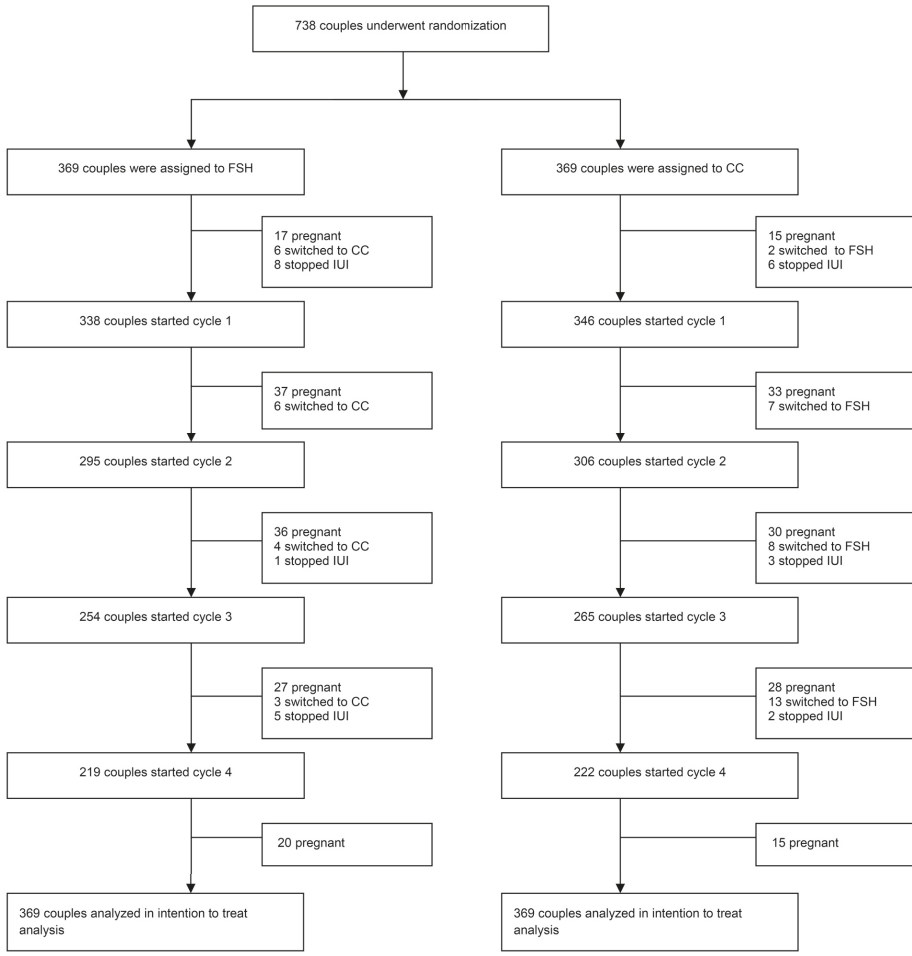
Between July 2013 and March 2016, we recruited 738 couples in 24 fertility clinics participating in the Dutch Consortium for Healthcare Evaluation in Obstetrics and Gynaecology (https://zorgevaluatienederland.nl/associations/1). A total of 369 couples were allocated to ovarian stimulation with FSH and 369 couples to ovarian stimulation with CC. The baseline characteristics were well balanced between couples that were randomized to FSH and those to CC (Table I). In the FSH treatment arm 338 couples received the allocated intervention and in the CC-treatment arm 346 couples (Fig. 1).

Table 1. Baseline characteristics of the participating couples*

Characteristics	FSH (*n*=369)	Clomiphene citrate (*n*=369)
Mean female age (years)	33·1 ±5·6	33·1 ± 4·6
Primary subfertility	273 (74)	268 (73)
Diagnosis of subfertility		
One-sided tubal pathology	28 (8)	38 (10)
Mild male subfertility	16 (4)	14 (4)
Median duration of subfertility (months)	24·0 (19·0 - 33·0)	24·0 (19·0 - 32·0)
Current smoking status	61 (17)	55 (15)
Mean body mass index in kg/m3	24·2 ±4·5	23·8 ± 3.9
Median total motile sperm count (x10^6)	48·0 (22·0 – 96·8)	58·4 (25·9 – 118·0)

* Data are n (%), mean (SD) or median (quartiles). There were no significant differences (P<0·05) between the two groups in any of the baseline characteristics.

Figure 1. Study flowchart



FSH=follicle stimulating hormone, CC=clomiphene citrate

Pregnancy outcomes are presented in Table II. Within the 6 months treatment horizon, there were 113 ongoing pregnancies (31%) in the FSH treatment arm and 97 ongoing pregnancies (26%) in the CC-treatment arm (RR 1.16, 95% CI: 0.93 to 1.47). The absolute risk difference for FSH compared to CC was 0.04 with a 95% CI of −0.02 to 0.11. In the per protocol analysis, there were 82 (25%) ongoing pregnancies in the FSH treatment arm and 70 ongoing pregnancies (21%) in the CC-treatment arm (RR 1.19, 95% CI: 0.90 to 1.57). In the FSH treatment arm, 17 women conceived naturally before they could start with IUI and nine women in between treatment cycles. In the CC-treatment arm, 15 women conceived naturally before they could start with IUI and seven women in between treatment cycles.

Table 2. Pregnancy outcomes per woman randomised*

	FSH (*n*=369)	Clomiphene citrate (*n*=369)	Relative Risk (95% CI)
Ongoing pregnancy	113 (31)	97 (26)	1·16 (0·93-1·47)
Multiple pregnancy	5 (1)	8 (2)	0·63 (0·21-1·89)
Live birth	105 (28)	92 (25)	1·14 (0·90-1·45)
Clinical pregnancy	115 (32)	101 (27)	1·14 (0·91-1·43)
Miscarriage	32 (9)	31 (8)	1·03 (0·64-1.66)
Ectopic pregnancy	2 (1)	3 (1)	0·80 (0·38-1·64)

* n (%)

The number of twin pregnancies was 5 (1.4%) in the FSH treatment arm and 8 (2.2%) in the CC-treatment arm (RR 0.63, 95% CI: 0.21 to 1.89, absolute rate difference [ARD] = −0.01, 95% CI: −0.03 to 0.01). There were no higher order multiple pregnancies. The number of live births was 105 (28%) in the FSH treatment arm and 92 (25%) in the CC-treatment arm (RR 1.14, 95% CI: 0.90 to 1.45).

Ovarian stimulation outcomes are shown in Table III. There was no difference in the cancellation rate due to the development of more than three dominant follicles between ovarian stimulation with FSH and ovarian stimulation with CC (FSH n = 115, CC n = 101, RR 1.06, 95% CI: 0.91 to 1.23). Other reasons for cycle cancellation were impaired folliculogenesis (FSH n = 32, CC n = 39), personal circumstances (FSH n = 9, CC n = 11) and other medical reasons (FSH n = 9, CC n = 2). There were slightly more cycles with monofollicular growth in ovarian stimulation with FSH compared to ovarian stimulation with CC (RR 1.12, 95% CI: 0.99 to 1.27).

Table 3. Ovarian stimulation outcomes on a cycle level*

	FSH (*n*=1162)	Clomiphene citrate (*n*=1212)	Relative Risk (95% CI)	*p*
Mean total dosage ovarian stimulation per cycle†	586 IU (328·9)	406 mg (423·1)	-	-
Mean duration of stimulation (days)†	8·1 (3·18)	4·9 (3·74)	-	-
Mean number of follicles ≥14 mm at day of ovulation triggering	1·8 (1·43)	1·9 (1·11)	-	0·52
Cycles with monofollicular growth	352 (30)	328 (27)	1·12 (0·99-1·27)	-
Cancellation rate	165 (14)	153 (13)	1·12 (0·92 – 1·38)	-
due to multifollicular growth	115 (70)	101 (66)	1·06 (0·91 – 1·23)	-

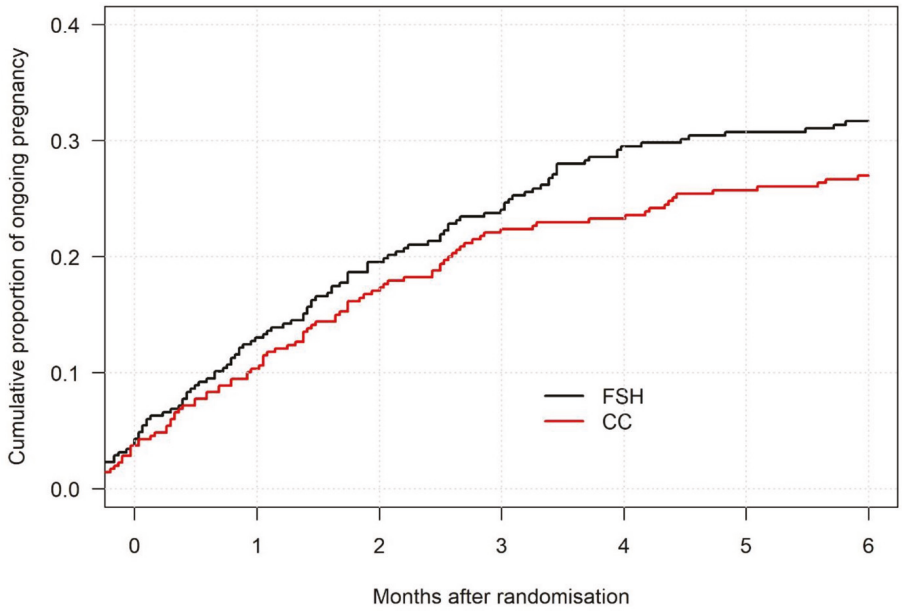
* Data are n (%), mean (SD)

† No p value was calculated since these outcomes are related to the type of ovarian stimulation

The multiple pregnancy rate was 0.2% after one dominant follicle and 0.7% after two dominant follicles (OR 3.3, 95% CI: 0.7 to 16.5), while it increased to 1.8% following three dominant follicles (OR 8.0 compared to one dominant follicle, 95% CI: 1.5 to 41.6)

In the intention to treat analysis, there was no difference in time to ongoing pregnancy between ovarian stimulation in the FSH treatment arm (P = 0.30) (Fig. 2). Likewise there was no difference in time to ongoing pregnancy between ovarian stimulation with FSH in the per protocol analysis (P = 0.30) (Fig. 3).

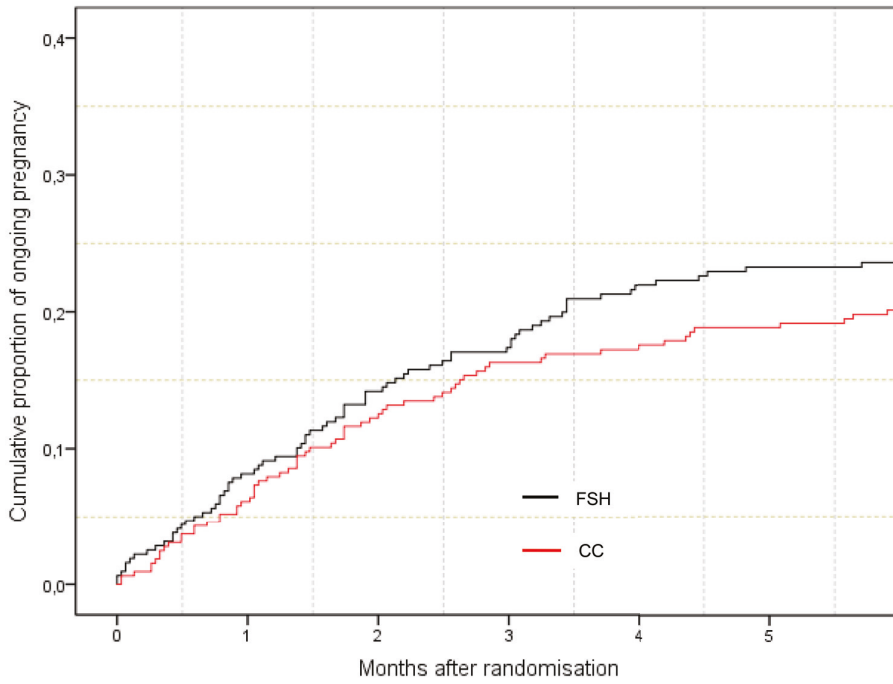
Figure 2. Time to ongoing pregnancy



FSH=follicle stimulating hormone, CC=clomiphene citrate

Numbers at risk							
	0	1	2	3	4	5	6
FSH	369	338	291	274	254	247	244
CC	369	346	301	284	272	264	262

Figure 3. Time to ongoing pregnancy – per protocol analysis



FSH=follicle stimulating hormone, CC=clomiphene citrate

Numbers at risk							
	0	1	2	3	4	5	6
FSH	333	333	280	262	242	236	233
CC	344	344	293	276	264	256	254

DISCUSSION

In this multicenter, non-blinded, randomized controlled superiority trial, we found no statistically significant difference between FSH and CC in couples with unexplained subfertility undergoing IUI with ovarian stimulation in a regimen of strict cancellation criteria, in terms of ongoing pregnancies, and a low multiple pregnancy rate. Our cumulative ongoing pregnancy rate of around 30% after four cycles of IUI within 6 months is comparable to the rates reported in a previous study, but we were able to reduce the high multiple pregnancy rate of 32% described in that study to 4% per cycle, which can be translated to a reduction of 11% to1% per woman (Diamond et al., 2015).

We feel that our findings are of importance, since IUI with ovarian stimulation is -as a first line treatment for couples with unexplained subfertility, applied worldwide on a large scale, as it is considered to be effective, but less invasive, less burdensome and less costly compared to IVF (Bensdorp et al., 2015). The new stimulation regimen described here can reduce the number of multiple pregnancies to such low levels that IUI with ovarian stimulation can now be regarded as a safe treatment if strict cancellation criteria are met.

The strength of this study is that, in our opinion, we had adequate power to show that there is no statistically significant difference between FSH and CC in IUI in terms of cumulative ongoing pregnancy rates ([ARD] = 0.04, 95% CI: −0.02 to 0.11), with both strategies leading to very low multiple pregnancy rates when adhering to strict cancellation criteria. This confirms the previous findings of the smaller study also aiming to reduce multiple pregnancy rate by means of adherence to strict cancellation criteria (Dankert et al., 2007) Our per protocol analysis showed the same results as the intention to treat analysis, suggesting that a switch in treatment did not affect the cumulative ongoing pregnancy rates, thereby underpinning the robustness of the data. We provided cumulative pregnancy outcomes because they give insight in the actual way of conceiving and represent true life. We were thus able to detect that as many as 48 (23%) ongoing pregnancies were conceived without medical assistance; 32 couples conceived before the start of IUI and 16 couples conceived in between treatment cycles. This again emphasizes that some of these couples, even though their prognosis of a natural conception was low and even though they were undergoing treatment, still manage to become pregnant in cycles without or in between treatment. This is important data to share with the couples in counselling.

Although our study is replication research, replication studies are fundamental in establishing progress, as they provide a more respectable basis of knowledge (Ioannidis, 2013). Pooling the data of our study and those of the smaller similar study, we find an ongoing pregnancy rate of 30% for FSH and of 27% for CC in IUI (RR 1.13, 95% CI: 0.91 to 1.40) and a multiple pregnancy rate of 1% for FSH and of 2% for CC in IUI (RR 1.65, 95% CI: 0.60 to 4.50) (Dankert et al., 2007).

Several limitations also need mentioning. According to ESHRE guidelines, live birth rate should be the primary outcome and we chose ongoing pregnancy as such. This is because ongoing pregnancy is seen as a valid and cost effective outcome measure of effectiveness (Clarke et al., 2010; Braakhekke et al., 2014). Nevertheless, we also do report on live birth rate. We were not able to blind this study due to the type of interventions. We consider it unlikely that this has introduced performance bias, since pregnancy outcomes are objective outcome measures. Another potential limitation of this study is that we based our sample size calculation on a 10% difference in ongoing pregnancy rate between the two stimulation agents. We can thus not rule out smaller differences. Future studies should thus be designed with a larger sample size to prove or reject any smaller difference.

Our results are widely generalizable, since the baseline characteristics of our patient population were similar to those reported in other international studies on IUI for unexplained subfertility (Cantineau and Cohlen, 2007; Bensdorp et al., 2015; Peeraer et al., 2015). As far as BMI is concerned, the mean BMI in studies from Europe are lower than those in the USA (Diamond et al., 2015).

With our regimen, the core of which is adherence to strict cancellation criteria, we were able to yield an average of two dominant follicles in both treatment arms. The strong association between an increase in the number of dominant follicles and multiple pregnancy, provide the rationale for this type of ovarian stimulation practice and explains its good safety profile (Rumste van et al., 2008). The question then rises whether ovarian stimulation with FSH or with CC should be the regimen of choice. A formal cost-effectiveness analysis will answer this question. Another strategy suggested to avoid multiple pregnancies in IUI for unexplained subfertility, has been selective ultrasound guided follicle aspiration prior to IUI when more than three dominant follicles develop (Stoop et al.,2010; Peeraer et al.,2015). Although this strategy has indeed been proven to be effective in reducing multiple pregnancies, it has never been compared to a strategy with adherence to strict cancellation criteria with respect to preference, burden and costs. Since patient care involves more domains than effectiveness, data are currently insufficient to advise this aspiration approach (Dancet et al.,2014).

Diamond et al. compared Letrozole to FSH and CC with multiple pregnancy as the primary outcome. The administration of Letrozole resulted in a similar low multiple pregnancy rate when compared to CC, but at the cost of live birth rates when compared to FSH. The live birth rates were 32% after FSH, 23% after CC and 19% after Letrozole (Diamond et al., 2015). At present, we cannot draw any firm conclusions on the effectiveness of Letrozole as a stimulation regimen in IUI for unexplained subfertility. Further studies are needed to investigate whether Letrozole can be considered in IUI for unexplained subfertility.

The discussion on single embryo transfer in IVF to reduce multiple pregnancies with its inherent risks for the mother and the offspring has taken years, when finally, single embryo transfer was successfully implemented and even today embryo transfer of more than one embryo is still common practice (Land and Evers, 2003, 2004; van Montfoort et al. 2005). Our study provides the protocol to also reduce multiple pregnancies in IUI with ovarian stimulation. Hopefully, this protocol will soon be implemented in clinical practice, regardless of the setting in which reproductive services are provided. In conclusion, we have shown that there is no statistically significant difference between an ovarian stimulation regimen with FSH compared to CC and adherence to strict cancellation criteria in couples with unexplained subfertility undergoing IUI in terms of ongoing pregnancies, live births and time to pregnancy, while yielding similar and low multiple pregnancy rates.

SUPPLEMENTARY DATA

Supplementary data are available at Human Reproduction online.

REFERENCES

Bensdorp AJ, Tjon-Kon-Fat RI, Bossuyt PM, Koks CA, Oosterhuis GJ, Hoek A, Hompes PG, Broekmans FJ, Verhoeve HR, de Bruin JP et al. Prevention of multiple pregnancies in couples with unexplained or mild male subfertility: randomised controlled trial of in vitro fertilisation with single embryo transfer or in vitro fertilisation in modified natural cycle compared with intrauterine insemination with controlled ovarian hyper- stimulation. Br Med J 2015;350:g7771.

Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infer- tility prevalence and treatment-seeking: potential need and demand for infertility medical care. Hum Reprod 2007;6:1506–1512.

Braakhekke M, Kamphuis EI, Dancet EA, Mol F, van der Veen F, Mol BW. Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper. Fertil Steril 2014;5:1203–1204.

Calhaz-Jorge C, De Geyter C, Kupka M, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C, Goossens V. European IVF-monitoring Consortium (EIM); European Society of Human Reproduction and Embryology (ESHRE), Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod 2017; 10:1957–1973.

Cantineau AEP, Cohlen BJ. Ovarian stimulation protocols (anti-oestrogens, gonadotrophins with and without GnRH agonists/antagonists) for intra- uterine insemination (IUI) in women with subfertility. Cochrane Database Syst Rev 2007.

Clarke JF, van Rumste MM, Farquhar CM, Johnson NP, Mol BW, Herbison P. Measuring outcomes in fertility trials: can we rely on clinical pregnancy rates? Fertil Steril 2010;5:1647–1651.

Dancet EA, D'Hooghe TM, van der Veen F, Bossuyt P, Sermeus W, Mol BW, Repping S. 'Patient-centered fertility treatment': what is required? Fertil Steril 2014;4:924–926.

Danhof NA, van Wely M, Koks C, Gianotten J, de Bruin JP, Cohlen BJ, van der Ham DP, Klijn NF, van Hooff MHA, Broekmans FJM et al. The

SUPER study: protocol for a randomised controlled trial comparing follicle-stimulating hormone and clomiphene citrate for ovarian stimula- tion in intrauterine insemination. BMJ Open 2017;5:e015680.

Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H, van Dop PA. A randomized clinical trial of clomifene cit- rate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. Hum Reprod. 2007;3:792–797.

Diamond MP, Richar SL, Coutifaris C, Alvero R, Robinson R, Casson P, Christman GM, Ager J, Huang H, Hansen KR et al. Letrozole, gonado- tropin, or clomiphene for unexplained infertility. N Engl J Med 2015;13: 1230–1240.

Guzick DS, Carson SA, Coutifaris C, Overstreet JW, Factor-Litvak P, Steinkampf MP, Hill JA, Mastroianni L, Buster JE, Nakajima ST et al. Efficacy of superovulation and intrauterine insemination in the treatment of infertility. National Cooperative Reproductive Medicine Network. N Engl J Med 1999;3:177–183.

Hunault CC, Habbema JDF, Eijkemans MJC, Collins JA, Evers JLH, te Velde ER. Two new prediction rules for spontaneous pregnancy leading to live birth among subfertile couples, based on the synthesis of three previous models. Hum Reprod 2004;19:2019–2026.

Consortium of the Dutch Society for Obstetrics and Gynaecology (inter- net). https:// zorgevaluatienederland.nl/associations/1 (9 May 9 2018, date last accessed).

Ioannidis JP. This I believe in genetics: discovery can be a nuisance, replica- tion is science, implementation matters. Front Genet 2013;4:33.

Lachin JM. Statistical considerations in the intent-to-treat principle. Control Clin Trials 2000;21:167–189.

Land JA, Evers JL. Risks and complications in assisted reproduction tech- niques: Report of an ESHRE consensus meeting. Hum Reprod 2003;2: 455–457.

Land JA, Evers JL. What is the most relevant standard of success in assisted reproduction? Defining outcome in ART: a Gordian knot of safety, effi- cacy and quality. Hum Reprod 2004;5:1046–1048.

National Institute for Health and Care Excellence. Fertility: for people with fer- tility problems. NICE Clinical Guideline. https://www.nice.org.uk/guidance/

4.

Intrauterine insemination for unexplained infertility – a network meta-analysis

authors:

N.A. Danhof, R. Wang, M. van Wely, F. van der Veen, B.W.J. Mol, M. H. Mochtar

Human Reproduction Update 2020;26(1):1-15

ABSTRACT

Background: IUI for unexplained infertility can be performed in a natural cycle or in combination with ovarian stimulation. A disadvantage of ovarian stimulation is an increased risk of multiple pregnancies with its inherent maternal and neonatal complication risks. Stimulation agents for ovarian stimulation are clomiphene citrate (CC), Letrozole or gonadotrophins. Although studies have compared two or three of these drugs to each other in IUI, they have never been compared to one another in one analysis.

Objective and rationale: The objective of this network meta-analysis, was to compare the effectiveness and safety of IUI with CC, Letrozole or gonadotrophins with each other and with natural cycle IUI.

Search methods: We searched PubMed, MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, CENTRAL, Clinical Trial Registration Database indexed up to 16-08-2018. We included randomized controlled trials that compared a stimulation regimen with CC, Letrozole or gonadotrophins to each other or to natural cycle IUI among couples with unexplained infertility. We performed the network meta-analysis within a multivariate random effects model.

Outcomes: We identified 26 studies reporting on 5316 women. The relative risk on live birth/ongoing pregnancy rates comparing IUI with CC to natural cycle IUI was 1.05(95%CI0.63-1.77, low quality of evidence), comparing IUI with Letrozole to natural cycle IUI was 1.15(95%CI0.63-2.08, low quality of evidence) and comparing IUI with gonadotrophins to natural cycle IUI was 1.46(95%CI0.92-2.30, low quality of evidence). The relative risk on live birth/ongoing pregnancy rates comparing gonadotrophins to CC was 1.39 (95% CI 1.09-1.76, moderate quality of evidence), comparing Letrozole to CC was 1.09 (95% CI 0.76-1.57, moderate quality of evidence) and comparing Letrozole to gonadotrophins was 0.79 (95% CI 0.54-1.15, moderate quality of evidence). We did not perform network meta-analysis on multiple pregnancy due to high inconsistency. Pairwise meta-analyses showed a relative risk on multiple pregnancy rates of 9.11 (95% CI 1.18-70.32) comparing IUI with gonadotrophins to natural cycle IUI. There was no data available on multiple pregnancy rates following IUI with CC or Letrozole compared to natural cycle IUI. The relative risk on multiple pregnancy rates comparing gonadotrophins to CC was 1.42 (95% CI 0.68-2.97), comparing Letrozole to CC was 0.97 (95% CI 0.47-2.01) and comparing Letrozole to gonadotrophins was 0.29 (95% CI 0.14-0.58).

In a meta-analysis among studies with adherence to strict cancellation criteria, the relative risk of live births/ongoing pregnancy rates comparing gonadotrophins to CC was 1.20 (95% CI 0.95-1.51) and the relative risk on multiple pregnancy rates comparing gonadotropins to CC was 0.80 (95% CI 0.38-1.68).

Wider implications: Based on low to moderate quality of evidence in this network meta-analysis, IUI with gonadotrophins ranked highest on live birth/ongoing pregnancy rates, but women undergoing this treatment protocol were also at risk for multiple pregnancies with high complication rates. IUI regimens with adherence to strict cancellation criteria lead to an acceptable multiple pregnancy rate without compromising the effectiveness. Within a protocol with adherence to strict cancellation criteria, gonadotrophins seem to improve live birth/ongoing pregnancy rates compared to CC. We, therefore, suggest performing IUI with ovarian stimulation with gonadotrophins within a protocol with strict cancellation criteria. Obviously, this ignores the impact of costs and patients preference.

INTRODUCTION

IUI is a first line treatment for unexplained infertility and is widely performed in many countries. It can be performed in a natural cycle or in combination with ovarian stimulation. Ovarian stimulation in IUI aims to increase the number of dominant follicles per cycle, based upon the concept that this will increase pregnancy rates (Rumste van M.M.E. 2008). Although IUI is a simple non-invasive procedure compared to IVF, it is not without severe medical risk in view of the high multiple pregnancy rates (2006, Kim, Child et al. 2015, Calhaz-Jorge, De Geyter et al. 2017). Multiple pregnancies bear the risk of serious maternal and neonatal complications (Ombelet, Martens et al. 2006).

Options for ovarian stimulation are clomiphene citrate (CC), Letrozole or gonadotrophins. CC and Letrozole are given orally for 5 days starting from cycle days 2 to 5. Gonadotrophins are administered as a s.c. injection from cycle days 3, 4 or 5 until the ovulation trigger. In a Cochrane review, CC has been compared to gonadotrophins and Letrozole has been compared to CC. The pregnancy rates were significantly increased when using gonadotrophins compared to CC (seven studies, 556 women, odds ratio (OR) 1.8, 95% CI 1.2 to 2.7) (Cantineau 2007). Pregnancy rates were similar comparing CC to Letrozole (five studies, 313 women, OR 1.2 95% CI 0.64 to 2.1). The power of the Cantineau (2007) meta-analysis was insufficient to reach any firm conclusions on multiple pregnancies. In a later large multicentre trial, multiple pregnancy rates were 10% for gonadotrophins, including six triplets, and 3% for CC and 3% for Letrozole without triplets (gonadotrophins versus CC, OR 4.6, 95% CI 2.11-10.22, gonadotrophins versus Letrozole, OR 4.22, 95% CI 1.99 – 8.95, CC versus Letrozole, OR 0.91, 95% CI 0.35 – 2.38) (Diamond, et al. 2015).

In IUI with ovarian stimulation, CC, Letrozole and gonadotrophins have never been compared to one another in one analysis and have never been compared to natural cycle IUI. We therefore used network analysis to compare the effectiveness and safety of IUI with CC, Letrozole or gonadotrophins and natural cycle IUI.

METHODS

In this network meta-analysis we followed the PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions (Hutton, et al. 2015). We registered the protocol at Prospero with registration number CRD42018105820 (https://www.crd.york.ac.uk/prospero/).

Search and study selection

We systematically searched PubMed, MEDLINE, EMBASE, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials (CENTRAL), and the Clinical Trial Registration Database indexed from inception to 16-08-2018. The details of the search are displayed in Supplementary Data I. Two authors (ND and RW) examined the identified studies independently for compliance with the inclusion criteria and selected eligible studies. They resolved disagreements by discussion with a third author (MW) and documented the selection process with a 'PRISMA' flowchart.

We included randomized controlled trials (RCTs) published on IUI with ovarian stimulation comparing CC, Letrozole or gonadotrophins with any other drug or with natural cycle IUI among couples with unexplained infertility. We included studies that randomized per woman and per cycle: in case of the latter, we only included the first treatment cycle. We excluded studies comparing dosages of the same drug.

Outcomes

The primary effectiveness outcome was live birth/ongoing pregnancy rates per woman randomized. We primarily used live birth rate and in case live birth was not reported, we used ongoing pregnancy rate. Live birth was defined as the birth of a child showing signs of life born at ≥20 weeks gestation or weighing ≥400 g. Ongoing pregnancy was defined as a registered heartbeat at or beyond 12 weeks of gestation. The primary safety outcome was multiple pregnancy rate per woman randomized, defined as registered heartbeat of at least two foetuses at 12 weeks of gestation. A secondary outcome was the clinical pregnancy rate per woman randomized, defined as any registered foetal heartbeat on ultrasound.

Data collection and analysis

Two authors (NA and RW) independently extracted data from the included studies and assessed the risk of bias of the included studies with the risk of bias assessment tool of the Cochrane Collaboration. We resolved disagreements by discussion with a third author (MW). We assessed selection bias, performance bias, detection bias, attrition bias, reporting bias, and other bias. In case of studies with non-blinded outcome assessors, we scored detection bias as at low risk because the outcome measures are objective. We assessed the overall quality of evidence by using a web application, Confidence in Network Meta-Analysis (CINeMA), which is based on the Grading of Recommendations, Assessment, Development and Evaluations framework.

We used network plots to illustrate all available head-to-head comparisons in included RCTs (Chaimani, et al. 2013). We tested for inconsistency by a design-by-treatment interaction model (Higgins, et al. 2012). If no inconsistency was detected, we performed network meta-analysis within the multivariate random effects meta-analysis model and calculated relative risks (RR) with 95% CIs and absolute risks. As a network meta-analysis does not involve new randomisation, it relies on the transitivity assumption. This assumption requires that all interventions compared in a network meta-analysis are jointly randomizable, i.e., all interventions compared in a network meta-analysis should be clinically reasonable in a theoretical multi-arm RCT. In our case, as natural cycle IUI or IUI with CC, Letrozole and gonadotrophins are all existing interventions in couples with unexplained infertility, we considered the transitivity assumption valid. We also performed pairwise meta-analyses in a random effects model in Stata, if direct data were available.

We used the surface under the cumulative ranking (SUCRA) to provide a hierarchy of the included interventions and used the comparison adjusted funnel plot to evaluate small study effects (Salanti, Ades et al. 2011, Chaimani, Higgins et al. 2013). The higher the SUCRA value, the better the rank of the treatment. For example, if SUCRA for treatment A is 100%, all treatments are worse than treatment A (I.e. A is the best); if SUCRA for treatment A is 0%, no treatment is worse than A (I.e. A is the worst); if SUCRA for treatment A is 50%, half of the treatments are worse than treatment A (Salanti, Ades et al. 2011, Chaimani, Higgins et al. 2013). For statistical analysis we used STATA software (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC, USA) (IR 2015).

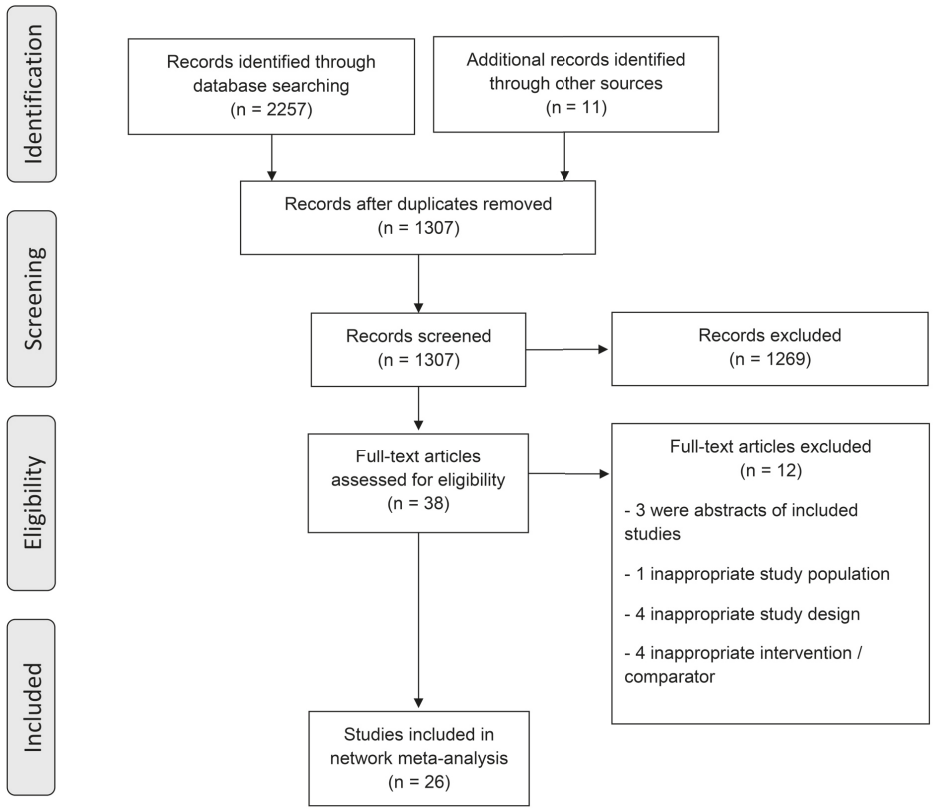
We performed a sensitivity analysis for studies with and without adherence to strict cancellation criteria with a maximum of three follicles of ≥14 mm per cycle. We also did a subgroup analysis in which we excluded studies that randomized per cycle.

RESULTS

Study selection

We identified 26 eligible studies reporting on 5316 couples with unexplained infertility. (Fig. 1) Excluded studies are presented in Supplementary Data II.

Figure 1. Prisma Flowchart



Study characteristics

The study characteristics are listed in Table I. We included 26 studies of 9 different countries. The mean age across studies ranged from 23 to 41 years and the mean duration of infertility from 2.5 to 5.9 years. Two studies were sponsored by pharmaceutical companies (Guzick, Carson et al. 1999, Dankert, Kremer et al. 2007). Five studies, totalling 750 women, compared IUI with ovarian stimulation to natural cycle IUI - two of which used gonadotrophins as stimulation agent (Guzick, Carson et al. 1999, Goverde, McDonnel et al. 2000) and three used CC as stimulation agent -(Martinez, Bernardus et al. 1990, Arici, Byrd et al. 1994, Leanza, Coco et al. 2014). In 10 studies totalling 2499 women, ovarian stimulation with gonadotrophins was compared to CC (Balasch, Ballesca et al. 1994, Kamel 1995, Nakajima, Smith et al. 1999, Dankert T 2007, Berker, Kahraman et al. 2011, Goldman, Thornton et al. 2014, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018). In nine studies totalling 2110 women, ovarian stimulation with Letrozole was compared to CC (Sammour 2001, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Badawy, Metwally et al. 2007, Fouda and Sayed 2011, Ibrahim, Moustafa et al. 2012, Diamond, Legro et al. 2015). In four studies totalling 830 women, ovarian stimulation with gonadotrophins was compared to Letrozole (Baysoy, Serdaroglu et al. 2006, Gregoriou, Vlahos et al. 2008, Diamond, Legro et al. 2015, Galal 2015). One study (900 women) had three intervention arms; CC, Letrozole and gonadotrophins (Diamond, Legro et al. 2015)) In one study, luteal support by vaginal progesterone was added in the gonadotrophins treatment arm (Galal 2015).

Three studies adhered to an IUI protocol with strict cancellation criteria (Dankert T 2007, Peeraer, Debrock et al. 2015, Danhof, et al. 2018). In one of these three studies selective ultrasound-guided follicular aspiration was performed or the cycle was cancelled; criteria for one or the other are not mentioned (Peeraer, et al. 2015).

Table 1. Characteristics of studies on IUI for unexplained infertility included in the analysis.

Study	Year	Country	Number of Patients	Intervention				Cancellation criteria	Outcome
				Natural cycle	CC	Letrozole	gonadotrophins		
Martinez et al.	1990	Netherlands	20	Yes	100 mg			No	Clinical pregnancy
Arici et al.	1994	USA	26	Yes	50 mg			No	Clinical pregnancy
Balasch et al.	1994	Spain	100		50 mg		75 IU	No	Clinical pregnancy, ongoing pregnancy
Kamel et al.	1995	Egypt	60		100 mg		75 IU	No	Clinical pregnancy
Guzick et al.	1999	USA	465	Yes			Dose unclear	No	Clinical pregnancy, live birth
Nakajima et al.	1999	Canada	22		Dose unclear		Dose unclear	No	Clinical pregnancy, multiple pregnancy
Goverde et al.	2000	Netherlands	171	Yes			Dose unclear	Max 3 follicles of ≥18 mm, or max 6 follicles ≥14 mm	Multiple pregnancy, live birth
Sammour et al.	2001	Canada	49		100 mg	2.5 mg		No	Clinical pregnancy
El Helw et al.	2002	Egypt	53		100 mg	20 mg		No	Clinical pregnancy
Fatemi et al.	2003	Belgium	15		100 mg	2.5 mg		No	Ongoing pregnancy
Al-Fozan et al.	2004	Canada	154		100 mg	7.5 mg		No	Ongoing pregnancy, multiple pregnancy
Ozmen et al.	2005	Turkey	43		100 mg	5 mg		No	Clinical pregnancy

Table 1. Continued

Study	Year	Country	Number of Patients	Intervention				Cancellation criteria	Outcome
				Natural cycle	CC	Letrozole	gonadotrophins		
Baysoy et al.	2006	Turkey	80			5 mg	75 IU (<30 years) and 150 IU (≥30 years)	No	Clinical pregnancy, multiple pregnancy
Dankert et al.	2007	Netherlands	138		100 mg		75 IU	Max 3 follicles of ≥14 mm	Clinical pregnancy, multiple pregnancy, live birth
Badawy et al.	2008	Egypt	412		100 mg	5 mg		No	Clinical pregnancy
Gregoriou et al.	2008	Greece	50			5 mg	150 IU	No	Clinical pregnancy, live birth
Berker et al.	2011	Turkey	189		100 mg		75 IU	Max 3 follicles of ≥17 mm or max 5 follicles ≥14 mm	Ongoing pregnancy, multiple pregnancy, live birth
Fouda et al.	2011	Egypt	214		100 mg	2.5 mg		No	Ongoing pregnancy, multiple pregnancy
Ibrahim et al.	2012	Egypt	270		100 mg	2.5 mg		No	Clinical pregnancy, multiple pregnancy
Goldman et al.	2014	USA	103		100 mg		300 IU	Max 6 follicles of >14 mm or poor ovarian response	Clinical pregnancy, multiple pregnancy, live birth
Leanza et al.	2014	Italy	68	Yes	50 mg				Clinical pregnancy

Table 1. Continued

Study	Year	Country	Number of Patients	Intervention				Cancellation criteria	Outcome
				Natural cycle	CC	Letrozole	gonadotrophins		
Diamond et al.	2015	USA	900		Dose unclear	Dose unclear	Dose unclear	No	Clinical pregnancy, multiple pregnancy, live birth
Erdem et al.	2015	Turkey	219		100 mg		75 IU	Max 4 follicles of ≥14 mm and/or max serum E2 levels of 1500 pg/ml	Clinical pregnancy, multiple pregnancy, live birth
Galal et al.	2015	Egypt	100			dose 2.5 mg increased daily by 2.5 mg for other 3 days	Dose unclear	No	Clinical pregnancy
Peeraer et al.	2015	Belgium	330		50 mg		37.5 or 75 IU	Max 3 follicles of ≥ 14 mm	Clinical pregnancy, multiple pregnancy, live birth
Danhof et al.	2018	Netherlands	738		100 mg		75 IU	Max 3 follicles of ≥14 mm or max 5 follicles ≥12 mm	Clinical pregnancy, multiple pregnancy, live birth

CC=clomiphene citrate

Quality assessment

The quality assessment of the included studies is shown in Figs 2 and 3. We scored nine studies as at unclear risk of random sequence selection bias (Martinez, Bernardus et al. 1990, Balasch, Ballesca et al. 1994, Kamel 1995, Nakajima, Smith et al. 1999, Sammour 2001, El Helw 2002, Ozmen 2005, Goldman, Thornton et al. 2014, Leanza, Coco et al. 2014) and 15 studies as at unclear risk of allocation concealment bias, because the procedures regarding allocation concealment were not properly described (Martinez, Bernardus et al. 1990, Arici, Byrd et al. 1994, Balasch, Ballesca et al. 1994, Kamel 1995, Nakajima, Smith et al. 1999, Sammour 2001, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Baysoy, Serdaroglu et al. 2006, Badawy, Metwally et al. 2007, Gregoriou, Vlahos et al. 2008, Leanza, Coco et al. 2014, Diamond, Legro et al. 2015, Galal 2015). We scored one study as at high risk of allocation concealment bias, because allocation was applied on day 3 of the treatment cycle in an open manner by only one instead of two investigators (Erdem, Abay et al. 2015). For blinding of participants and personnel, we scored 21 studies as at unclear risk because they were open label studies (Martinez, Bernardus et al. 1990, Arici, Byrd et al. 1994, Balasch, Ballesca et al. 1994, Kamel 1995, Guzick, Carson et al. 1999, Nakajima, Smith et al. 1999, Goverde, McDonnell et al. 2000, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Badawy, Metwally et al. 2007, Dankert T 2007, Gregoriou, Vlahos et al. 2008, Berker, Kahraman et al. 2011, Goldman, Thornton et al. 2014, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Galal 2015, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018). We scored the risk of bias for blinding of outcome assessment as at low risk for all studies, because of the objectivity of the outcome measure. For incomplete data outcome, we scored nine studies as at unclear risk (Arici, Byrd et al. 1994, Nakajima, Smith et al. 1999, Sammour 2001, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Baysoy, Serdaroglu et al. 2006, Leanza, Coco et al. 2014), because not enough information was provided to judge this and we scored one study as at high risk of bias because 10% of the included patients were excluded from the analysis (Kamel 1995). For selective reporting, we scored 13 studies as at unclear risk, because not enough information was provided on this topic to judge (Martinez, Bernardus et al. 1990, Arici, Byrd et al. 1994, Balasch, Ballesca et al. 1994, Kamel 1995, Guzick, Carson et al. 1999, Nakajima, Smith et al. 1999, Sammour 2001, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Ozmen 2005, Gregoriou, Vlahos et al. 2008, Leanza, Coco et al. 2014, Galal 2015). In 17 studies, we could not evaluate any other risks of bias due to lack of information (Martinez, Bernardus et al. 1990, Arici, Byrd et al. 1994, Balasch, Ballesca et al. 1994, Kamel 1995, Guzick, Carson et al. 1999, Nakajima, Smith et al. 1999, Sammour 2001, El Helw 2002, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Baysoy, Serdaroglu et al. 2006, Dankert T 2007, Gregoriou, Vlahos et al. 2008, Leanza, Coco et al. 2014, Galal 2015, Peeraer, Debrock et al. 2015). We scored one study as at high risk of observer bias because we found a subsequent abstract with different results (Badawy, Metwally et al. 2007).

Figure 2. Risk of Bias graph of the included studies according to the bias assessment tool of the Cochrane Collaboration

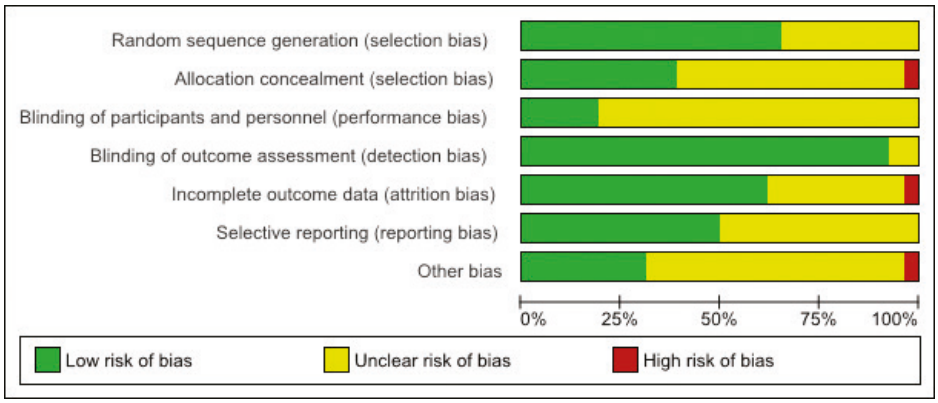
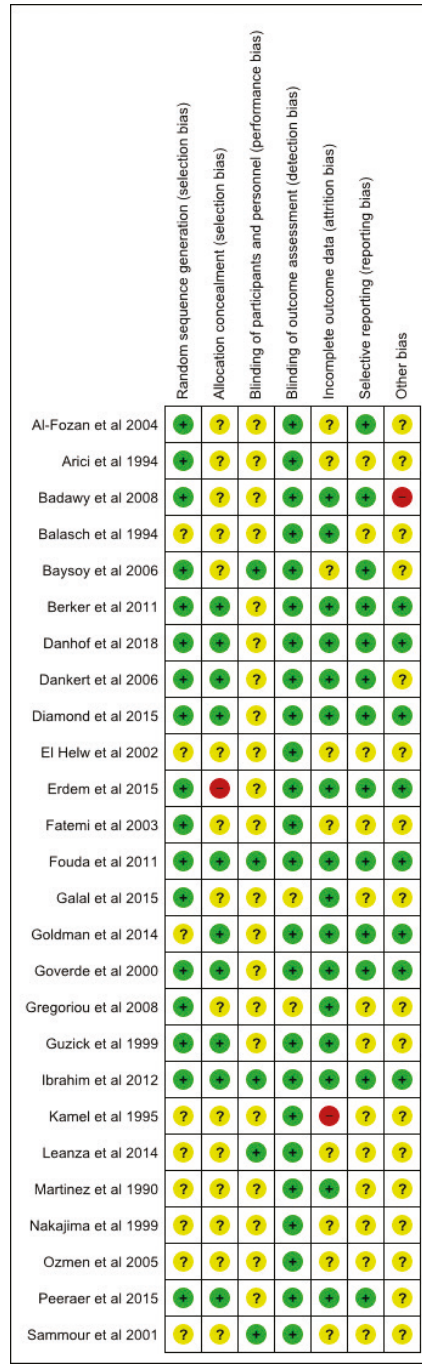


Figure 3. Risk of bias summary of the included studies according to the bias assessment tool of the Cochrane Collaboration



Network and pairwise meta-analyses

Live birth/ongoing pregnancy

Fourteen studies (4113 women) reported on live birth rate/ongoing pregnancy rates (Balasch, Ballesca et al. 1994, Guzick, Carson et al. 1999, Goverde, McDonnell et al. 2000, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Dankert T 2007, Gregoriou, Vlahos et al. 2008, Berker, Kahraman et al. 2011, Fouda and Sayed 2011, Goldman, Thornton et al. 2014, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018). The results of the network meta-analysis, pairwise meta-analysis and overall quality of evidence of live birth/ongoing pregnancy rates are shown in Table II. The results of the network meta-analysis are shown in Fig. 4 and also depicted in Fig. 5.The RR for live birth/ongoing pregnancy rates comparing CC to natural cycle IUI was 1.05 (95%CI 0.63 to 1.77, low quality of evidence). The RR for live birth/ongoing pregnancy rates comparing IUI with gonadotrophins to natural cycle IUI was 1.46 (95%CI 0.92 to 2.30, low quality of evidence) and comparing Letrozole to natural cycle IUI was 1.15 (95%CI 0.63 to 2.08, low quality of evidence). The network meta-analysis showed that the RR for live birth/ongoing pregnancy rates comparing gonadotrophins to CC was 1.39 (95%CI 1.09 to 1.76, moderate quality of evidence). The RR for live birth/ongoing pregnancy rates comparing Letrozole to CC was 1.09 (95%CI 0.76 to 1.57, moderate quality of evidence) and comparing Letrozole to gonadotropins was 0.79 (95%CI 0.54 to 1.15, moderate quality of evidence). The results from the pairwise meta-analysis were consistent with network meta-analysis. IUI with gonadotrophins was the most effective (SUCRA 93.8%), followed by Letrozole (SUCRA 49.5%), CC (SUCRA 30.7%) and natural cycle IUI (SUCRA 26.0%) (Fig. 5).

Table 2. Pairwise and network meta-analysis for live birth/ongoing pregnancy

Treatment comparison		Pairwise meta-analysis		Network meta-analysis	Overall certainty of evidence based on GRADE assessment
		No of studies	RR (95% CI)	RR (95% CI)	
CC		0	NA	1.05 (0.63 - 1.77)	Low[1,2]
Gonadotrophins	Natural cycle	2	1.46 (1.09 - 1.98)	1.46 (0.92 - 2.30)	Low[1,2]
Letrozole		0	NA	1.15 (0.63 - 2.08)	Low[1,2]
Gonadotrophins	CC	8	1.32 (1.06 - 1.64)	1.39 (1.09 - 1.76)	Moderate[3]
Letrozole		4	1.19 (0.70 - 2.02)	1.09 (0.76 - 1.57)	Moderate[2]
Letrozole	Gonadotrophins	2	0.59 (0.45 - 0.78)	0.79 (0.54 - 1.15)	Moderate[2]

Notes:
1. Downgraded by one level due to some concerns on the risk of bias.
2. Downgraded by one level due to some concerns on imprecision.
3. Downgraded by one level due to some concerns on heterogeneity.

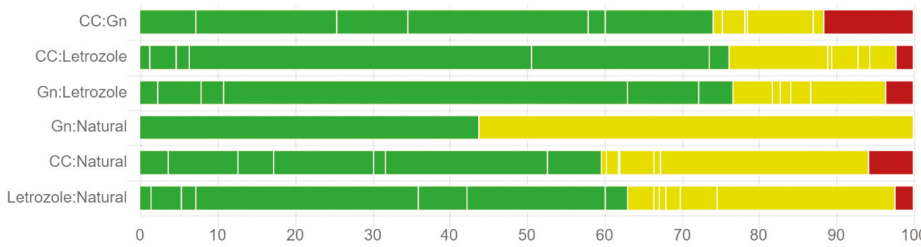


Figure 4. Network meta-analysis for live birth/ongoing pregnancy comparing a stimulation regimen with CC, Letrozole or gonadotrophins to each other or to natural cycle IUI, presented in relative risks (CC=clomiphene citrate, Gn=gonadotrophins)

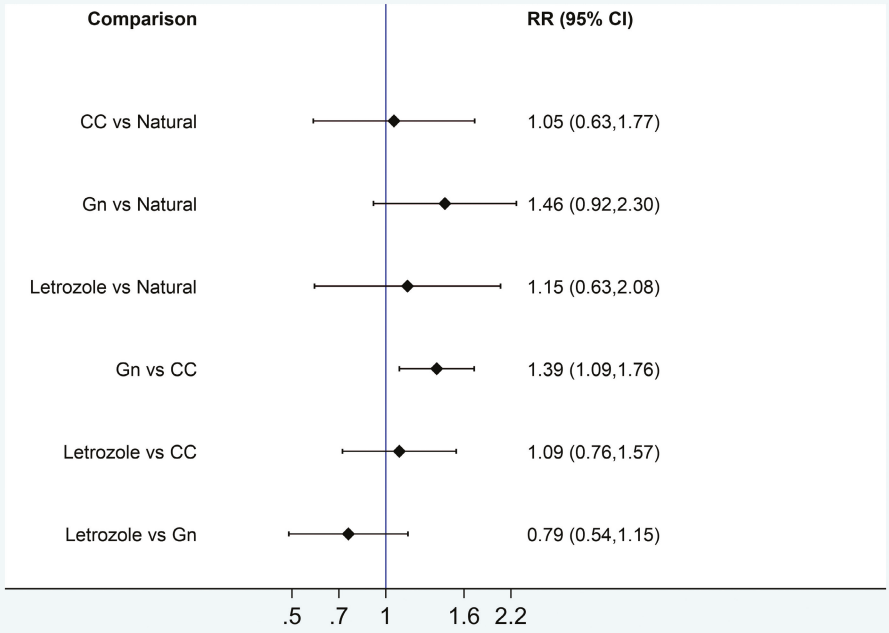
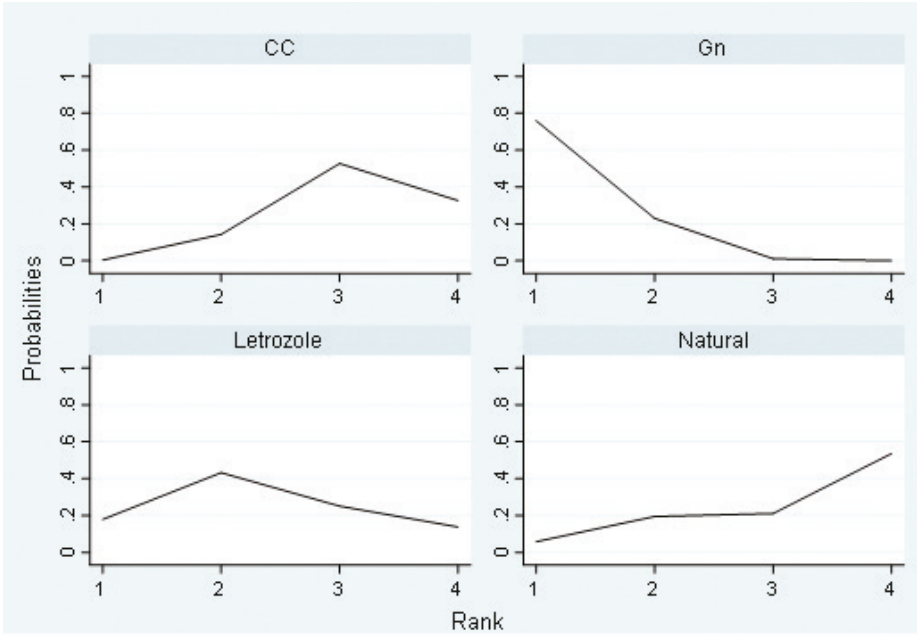


Figure 5. Network meta-analysis for clinical pregnancy comparing a stimulation regimen with CC, Letrozole or gonadotrophins to each other or to natural cycle IUI, presented in relative risks (CC=clomiphene citrate, Gn=gonadotrophins)



CC=Clomiphene citrate, GN=gonadotrophins

Multiple pregnancy

Thirteen studies (3855 women), reported on multiple pregnancy as an outcome (Nakajima, Smith et al. 1999, Goverde, McDonnell et al. 2000, Al-Fozan, Al-Khadouri et al. 2004, Baysoy, Serdaroglu et al. 2006, Dankert T 2007, Berker, Kahraman et al. 2011, Fouda and Sayed 2011, Ibrahim, Moustafa et al. 2012, Goldman, Thornton et al. 2014, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018). The global inconsistency test showed significant inconsistency (Chi^2=10.42, p =0.01). Therefore, we did not perform a network meta-analysis. The results of the pairwise meta-analysis are shown in Table III. The resulted presented in odds ratios are shown in supplementary data III.

Table 3. Pairwise meta-analysis for multiple pregnancy

Treatment comparison		Pairwise meta-analysis		Overall certainty of evidence based on GRADE assessment RR (95% CI)
		No of studies	RR (95% CI)	
CC		0	NA	-
Gonadotrophins	Natural cycle	1	9.11 (1.18 - 70.32)	Moderate[1]
Letrozole		0	NA	-
Gonadotrophins	CC	8	1.42 (0.68 - 2.97)	Low[2,3]
Letrozole		4	0.97 (0.47 - 2.01)	Moderate[1]
Letrozole	Gonadotrophins	2	0.29 (0.14 - 0.58)	Moderate[1]

RR: relative risk, GRADE: Grading of Recommendations, Assessment, Development and Evaluations, Gn: gonadotrophins

Notes:

1. Downgraded by one level due to some concerns on imprecision.
2. Downgraded by one level due to some concerns on the risk of bias.
3. Downgraded by one level due to some concerns on heterogeneity.

Clinical pregnancy

Twenty-one studies (4246 women) reported clinical pregnancy as an outcome (Arici, Byrd et al. 1994, Balasch, Ballesca et al. 1994, Kamel 1995, Guzick, Carson et al. 1999, Nakajima, Smith et al. 1999, Goverde, McDonnell et al. 2000, Sammour 2001, Fatemi, Kolibianakis et al. 2003, Al-Fozan, Al-Khadouri et al. 2004, Ozmen 2005, Badawy, Metwally et al. 2007, Gregoriou, Vlahos et al. 2008, Berker, Kahraman et al. 2011, Fouda and Sayed 2011, Goldman, Thornton et al. 2014, Leanza, Coco et al. 2014, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Galal 2015, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018). The results of the network meta-analysis, pairwise meta-analysis and overall quality of evidence on clinical pregnancy are shown in Table IV and Fig. 6. The RR for clinical pregnancy rates comparing ovarian stimulation with CC to natural cycle IUI was 1.89 (95%CI 1.14 to 3.12, low quality of evidence), comparing IUI with gonadotrophins to natural cycle IUI was 2.32 (95%CI 1.41 to 3.80, moderate quality of evidence) and comparing IUI with Letrozole to natural cycle IUI was 2.18 (95%CI 1.25 to 3.82, low quality of evidence). The RR for clinical pregnancy rates comparing gonadotrophins to CC was 1.23 (95%CI 1.00 to 1.51, low quality of evidence). The RR for clinical pregnancy rates comparing Letrozole to CC was 1.09 (95%CI 0.76 to 1.57). The RR for clinical pregnancy rates comparing Letrozole to gonadotrophins was 0.94 (95%CI 0.71 to 1.25, low quality of evidence). IUI with gonadotrophins was most likely to improve clinical pregnancy rates (SUCRA 87.8%), followed by Letrozole (SUCRA 72.4%), CC (SUCRA 39.5%) and Natural cycle IUI (SUCRA 0.3%).

Table 4. Pairwise and network meta-analysis for clinical pregnancy

Treatment comparison		Pairwise meta-analysis		Network meta-analysis	Overall certainty of evidence based on GRADE assessment
		No of studies	RR (95% CI)	RR (95% CI)	
CC		3	3.08 (1.41 - 6.73)	1.89 (1.14 - 3.12)	Low[1,2]
Gonadotrophins	Natural cycle	1	1.86 (1.34 - 2.58)	2.32 (1.42 - 3.80)	Moderate[1]
Letrozole		NA	NA	2.18 (1.25 - 3.82)	Low[1,2]
Gonadotrophins	CC	9	1.19 (0.99 - 1.43)	1.23 (1.00 - 1.51)	Low[1,2]
Letrozole		7	1.27 (0.91 - 1.78)	1.16 (0.90 - 1.49)	Low[1,3]
Letrozole	Gonadotrophins	4	0.67 (0.53 - 0.85)	0.94 (0.71 - 1.25)	Low[1,3]

RR: relative risk, GRADE: Grading of Recommendations, Assessment, Development and Evaluations, Gn: gonadotrophins
1. Downgraded by one level due to some concerns on the risk of bias.
2. Downgraded by one level due to some concerns on inconsistency (heterogeneity or incoherence).
3. Downgraded by one level due to some concerns on imprecision.

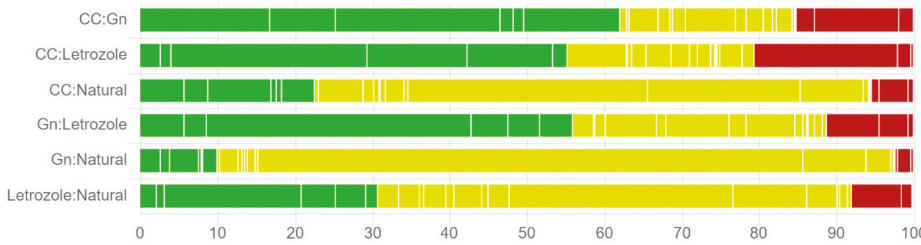
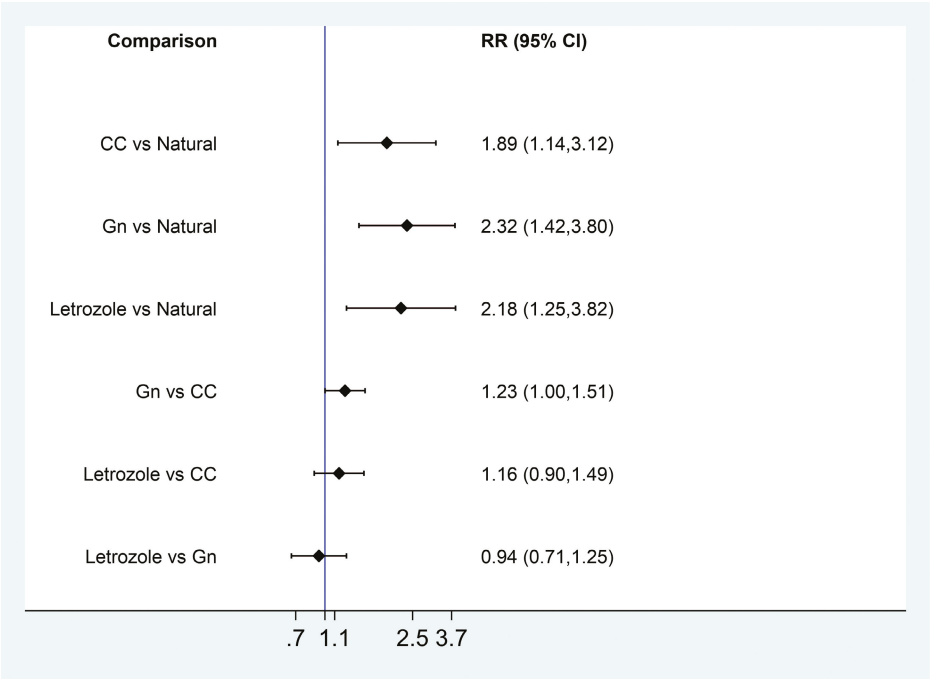


Figure 6. Subgroup analysis for live birth/ongoing pregnancy in studies with and without adherence to strict cancellation criteria, ie cycle cancellation when > 3 follicles developed, presented in relative risks (CC=clomiphene citrate, Gn=gonadotrophins)



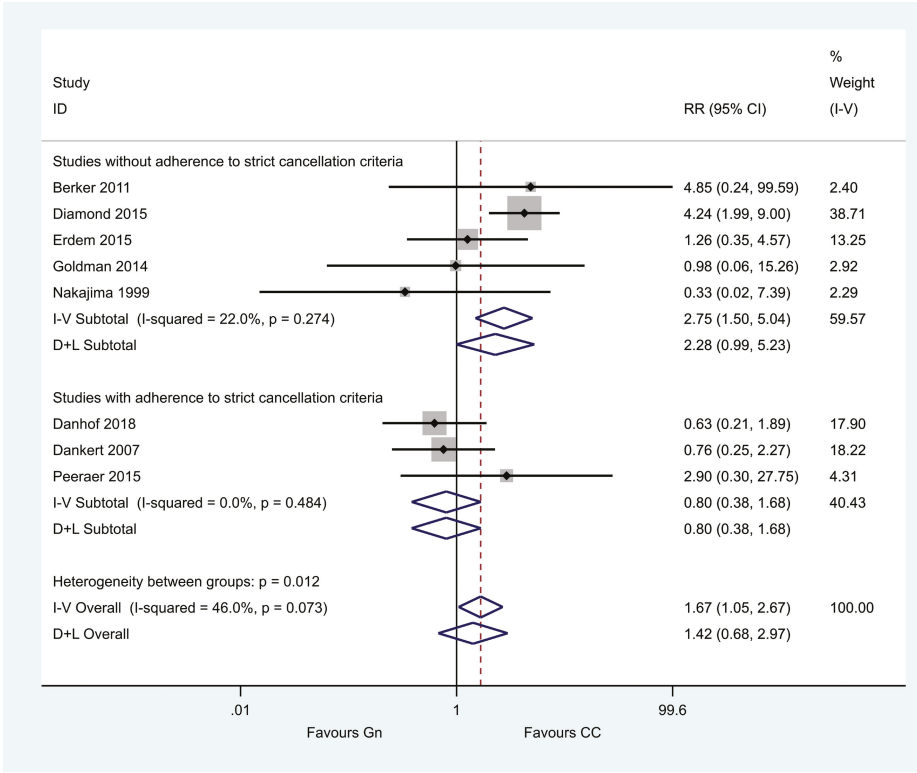
Subgroup analysis of studies without and with adherence to strict cancellation criteria

Live birth/ ongoing pregnancy

In five studies reporting on 1539 women, IUI was performed without adherence to strict cancellation criteria.(Berker, Kahraman et al. 2011, Diamond, Legro et al. 2015, Erdem, Abay et al. 2015, Goldman, Thornton et al. 2014, Nakajima, Smith et al. 1999).In these studies, the RR for live births/ongoing pregnancy rates comparing gonadotrophins to CC was 1.46 (95%CI 1.01 to 2.12, random-effect model, I2=60.8%).

In three studies reporting on 1206 women, IUI was performed with adherence to strict cancellation criteria (Dankert T 2007, Peeraer, Debrock et al. 2015, Danhof, van Wely et al. 2018) (Fig. 7). In these studies, the RR for live births/ongoing pregnancy rates comparing gonadotrophins to CC was 1.20 (95%CI 0.95 to 1.51, random-effect model, I2=16.1%). There was no statistically significant subgroup effect (p= 0.304).

Figure 7. Subgroup analysis for multiple pregnancy in studies with and without adherence to strict cancellation criteria, ie cycle cancellation when > 3 follicles developed, presented in relative risks (CC=clomiphene citrate, Gn=gonadotrophins)



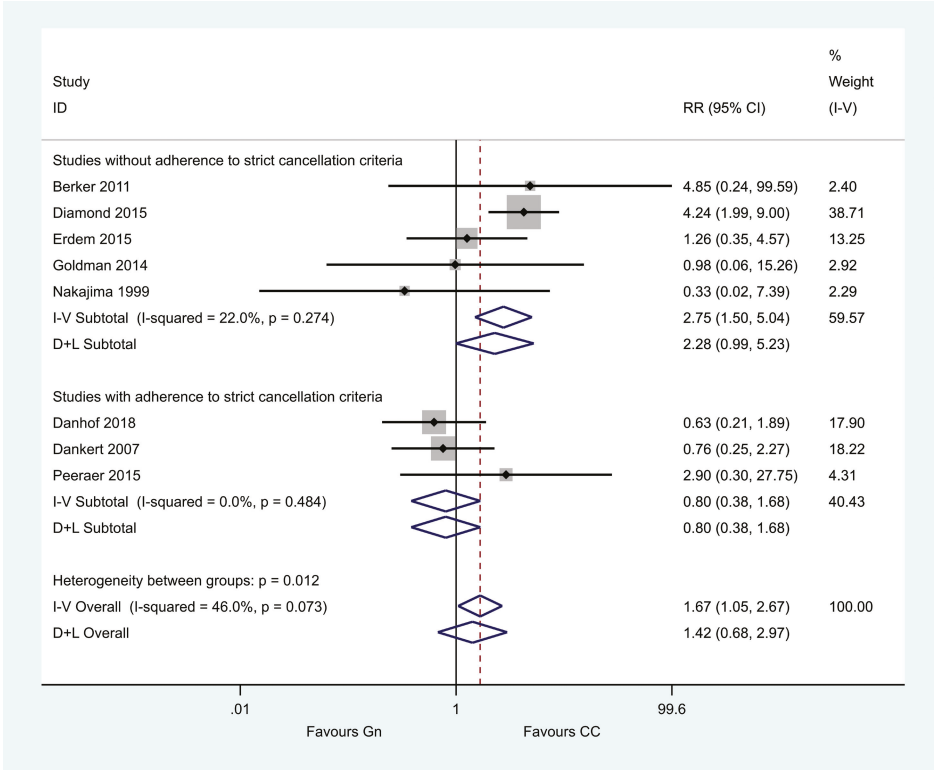
I-V: The inverse-variance fixed-effect method
D+L: The DerSimonian and Laird random-effect method
Favours Gn means less multiple pregnancies in the gonadotrophins group

Multiple pregnancy

In studies without adherence to strict cancellation criteria, the relative risk on multiple pregnancy rates comparing gonadotrophins to CC was 2.28 (95% CI 0.99 to 5.23, random-effect model, I^2=22.0%)

In studies with adherence to strict cancellation criteria, the relative risk on multiple pregnancy rates comparing gonadotrophins to CC was 0.80 (95% CI 0.38-1.68, I^2=0). There was a statistically significant subgroup effect (p= 0.012) (figure 8).

Figure 8. Sensitivity analysis on live birth/ongoing pregnancy for studies that randomized per woman (CC=clomiphene citrate, Gn=gonadotrophins)

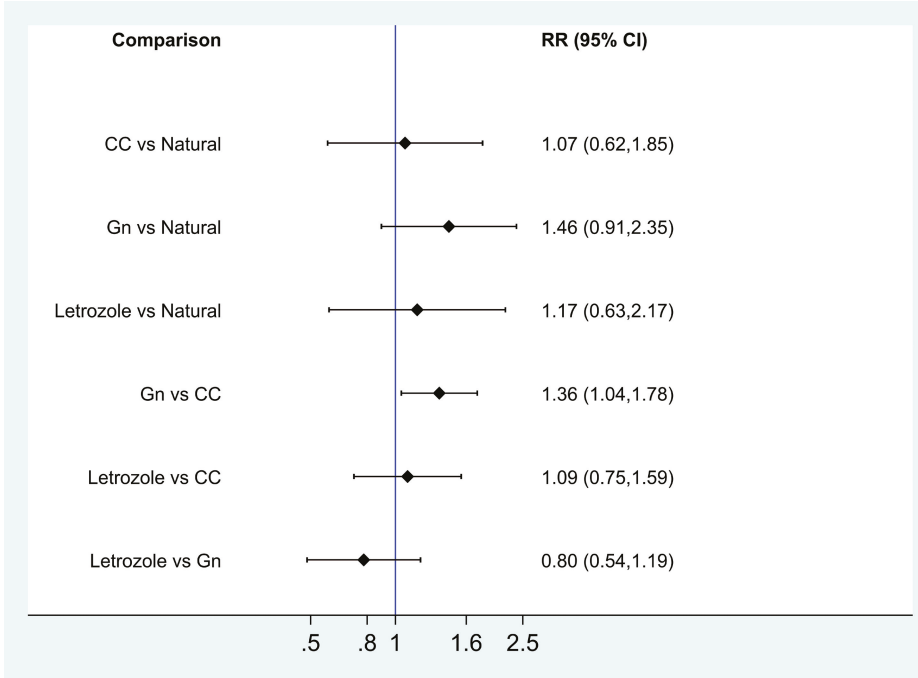


I-V: The inverse-variance fixed-effect method
D+L: The DerSimonian and Laird random-effect method
Favours Gn means less multiple pregnancies in the gonadotrophins group

Sensitivity analysis excluding studies that randomized per cycle

In this post-hoc sensitivity analysis, we excluded the study Peeraer 2015 as this study randomized cycles instead of women (Peeraer, Debrock et al. 2015). The sensitivity analysis was consistent with the main analysis on the primary effectiveness outcome (figure 9).

Figure 9. Sensitivity analysis on live birth/ongoing pregnancy excluding studies that randomized per cycle



DISCUSSION

Principal findings

In this network meta-analysis, we evaluated IUI with ovarian stimulation with CC, Letrozole or gonadotrophins and natural cycle IUI among couples with unexplained infertility. IUI with gonadotrophins ranked highest on live birth rate/ongoing pregnancy rates, whereas natural cycle IUI ranked the lowest on these outcomes. Multiple pregnancy rates were increased following IUI with gonadotrophins. Adhering to strict cancellation criteria reduced these rates.

Strengths and limitations

This study has several strengths and limitations that need to be considered. A first strength is that we reported on the most important outcome measures live birth and ongoing pregnancy.(Land and Evers 2003, Braakhekke, Kamphuis et al. 2014) Consequently, our statistical power was limited as not all studies reported on live birth rate or ongoing pregnancy rate. Therefore, we therefore also reported on clinical pregnancy.

Second, the tool of network meta-analysis gave us the opportunity to rank the treatment strategies in order of effectiveness, facilitating clinical decision making

As far as the limitations are concerned, we have to acknowledge that we were not able to perform network meta-analysis for the primary safety outcome multiple pregnancy rate because of high heterogeneity. We have endeavoured to present the evidence as complete as possible, since we reported a pairwise meta-analysis and a subgroup analysis including studies that performed IUI with and without adherence to strict cancellation criteria with multiple pregnancy as the primary safety outcome. Therefore, we feel that the evidence regarding to gonadotrophins and clomiphene citrate in relation to multiple pregnancies is robust. When interpreting the results, the overall quality of evidence is of paramount importance. The overall quality of evidence was downgraded to low or moderate, due to concerns on risk of bias, imprecision and heterogeneity. Another point to mention is that our reported outcome measures were per woman randomized. For live birth rate this is the only relevant outcome measure, but in multiple pregnancy this ensures a shortcoming, because it is partially a measure of success; one cannot have a multiple pregnancy without becoming pregnant. Nevertheless, we feel that this is the best possible approach, since all women randomized were at risk of a multiple pregnancy. A weakness by design is that we could not adjust for possible confounders such as female age, BMI, duration of infertility, primary versus secondary infertility or diagnosis of infertility, follicle size, cancellation criteria and endometrial thickness at day of ovulation triggering, as we had no access to individual patient data. An ongoing debate is obviously the definition of unexplained infertility which is not internationally established, which has led to different approaches among studies. For example, some studies included women with mild endometrioses, while others excluded women with endometriosis. Just as HPV infections, which can influence pregnancy rates following IUI, were not mentioned in the included studies. (Depuydt CE 2016, Depuydt CE 2019) Also the sperm quality differed among the included studies.

Comparison with other studies

Our study is the first network meta-analysis that compared IUI with ovarian stimulation with CC, Letrozole or gonadotrophins and natural cycle IUI to one another, including a subgroup analysis in which we analysed studies that performed IUI with adherence to strict cancellation criteria to avoid an increased risk of multiple pregnancies.

Explanation of results

Our data showed that protocols with adherence to strict cancellation criteria, are safe with regard to multiple pregnancies without significantly compromising on the live birth and ongoing pregnancy rates, proved by the test for subgroup effect, in which we did find a statistically significant effect of strict cancellation criteria on multiple pregnancy rates and no statistically significant effect of strict cancellation criteria on live birth/ongoing pregnancy rates. The data are easily explained based on the strong association between an increase in the number dominant follicles and an increase in multiple pregnancies.(Rumste van M.M.E. 2008) It is biologically plausible that ovarian stimulation with a maximum of three dominant follicles per cycle lead to similar pregnancy outcomes, despite the different endocrinological mechanisms by which CC,

Letrozole and gonadotrophins operate. These differences are no longer relevant if their impact on folliculogenesis is limited by active interventions like cancelling cycles above a specific cut off level.

Implications for clinical practice

In view of the data on effectiveness and safety shown in this study, we suggest performing IUI with ovarian stimulation with gonadotrophins within a protocol with strict cancellation criteria. Obviously, this ignores the impact of costs and patients preference.

Implications for further research

First, we suggest a cost-effectiveness analysing comparing IUI with CC, Letrozole, gonadotrophins and natural cycle IUI. Second, the choice for CC or Letrozole could be interesting from a patients point, ie oral tablets instead of injections. In our study, the evidence on the comparison between CC and Letrozole is limited due to heterogeneity. We therefore, suggest a new randomized controlled trial comparing CC to Letrozole within a protocol with adherence to strict cancellation criteria and if applicable, a cost-effectiveness analysis alongside the trial. Finally, we suggest an individual patient data analysis to investigate possible prognostic variables, such as female age, BMI, duration of infertility, primary versus secondary infertility or diagnosis of infertility, and to identify whether there are specific populations that benefit from a specific stimulation regimens in IUI for unexplained infertility, ie whether it is possible to treat an a more individual level.

CONCLUSION

Based on low to moderate quality of evidence in this network meta-analysis, IUI with gonadotrophins ranked highest on live birth rate/ongoing pregnancy rates, whereas natural cycle IUI ranked the lowest on these outcomes. Multiple pregnancy rates were increased following IUI with gonadotrophins. Adhering to strict cancellation criteria reduced these rates.

SUPPLEMENTARY DATA

Supplementary data are available online

ACKNOWLEDGEMENTS

We thank Marian Showell, information specialist of Cochrane Gynaecology and Fertility for performing the systematic search.

COMPETING INTERESTS

Prof dr BWJ Mol is supported by a NHMRC Practitioner Fellowship (GNT1082548) and reports consultancy for Merck, ObsEva and Guerbet.

FUNDING

This study is not externally funded.

REFERENCES

"CINeMA: Confidence in Network Meta-Analysis [Software]." 2017.

(2006). "Effectiveness and treatment for unexplained infertility." Fertil Steril 86(5 Suppl 1): S111-114.

Al-Fozan, H., M. Al-Khadouri, S. L. Tan and T. Tulandi (2004). "A randomized trial of letrozole versus clomiphene citrate in women undergoing superovulation." Fertility and Sterility 82(6): 1561-1563.

Arici, A., W. Byrd, K. Bradshaw, W. H. Kutteh, P. Marshburn and B. R. Carr (1994). "Evaluation of clomiphene citrate and human chorionic gonadotropin treatment: a prospective, randomized, crossover study during intrauterine insemination cycles." Fertil Steril 61(2): 314-318.

Badawy, A. M., M. Metwally and M. Fawzy (2007). "Randomized controlled trial of three doses of letrozole for ovulation induction in patients with unexplained infertility." Reproductive BioMedicine Online 14(5): no pagination.

Balasch, J., J. L. Ballesca, C. Pimentel, M. Creus, F. Fabregues and J. A. Vanrell (1994). "Late low-dose pure follicle stimulating hormone for ovarian stimulation in intra-uterine insemination cycles." Human Reproduction 9(10): 1863-1866.

Bayar U, Tanriverdi HA, Barut A, Ayoğlu F, Ozcan O, Kaya E. Letrozole vs. clomiphene citrate in patients with ovulatory infertility. Fertil Steril. 2006 Apr;85(4):1045-8.

Baysoy, A., H. Serdaroglu, H. Jamal, E. Karatekeli, H. Ozornek and E. Attar (2006). "Letrozole versus human menopausal gonadotrophin in women undergoing intrauterine insemination." Reproductive BioMedicine Online 13(2): 208-212.

Berker, B., K. Kahraman, S. Taskin, Y. E. Sukur, M. Sonmezer and C. S. Atabekoglu (2011). "Recombinant FSH versus clomiphene citrate for ovarian stimulation in couples with unexplained infertility and male subfertility undergoing intrauterine insemination: A randomized trial." Archives of Gynecology and Obstetrics 284(6): 1561-1566.

Braakhekke, M., E. I. Kamphuis, E. A. Dancet, F. Mol, F. van der Veen and B. W. Mol (2014). "Ongoing pregnancy qualifies best as the primary outcome measure of choice in trials in reproductive medicine: an opinion paper." Fertil Steril 101(5): 1203-1204.

Calhaz-Jorge, C., C. De Geyter, M. S. Kupka, J. de Mouzon, K. Erb, E. Mocanu, T. Motrenko, G. Scaravelli, C. Wyns and V. Goossens (2017). "Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE." Hum Reprod 32(10): 1957-1973.

Cantineau (2007). "Ovarian stimulation protocols (anti-oestrogens, gonadotrophins with and without GnRH agonists/antagonists) for intrauterine insemination (IUI) in women with subfertility. ." Cochrane Database Syst Rev.

Chaffkin LM, Nulsen JC, Luciano AA, Metzger DA. A comparative analysis of the cycle fecundity rates associated with combined human menopausal gonadotropin (hMG) and intrauterine insemination (IUI) versus either hMG or IUI alone. Fertil Steril. 1991 Feb;55(2):252-7.

Chaimani, A., J. P. Higgins, D. Mavridis, P. Spyridonos and G. Salanti (2013). "Graphical tools for network meta-analysis in STATA." PLoS One 8(10): e76654.

Danhof, N. A., M. van Wely, S. Repping, C. Koks, H. R. Verhoeve, J. P. de Bruin, M. F. G. Verberg, M. H. A. van Hooff, B. J. Cohlen, C. F. van Heteren, K. Fleischer, J. Gianotten, J. van Disseldorp, J. Visser, F. J. M. Broekmans, B. W. J. Mol, F. van der Veen, M. H. Mochtar and S. s. group (2018). "Follicle stimulating hormone versus clomiphene citrate in intrauterine insemination for unexplained subfertility: a randomized controlled trial." Human Reproduction: dey268-dey268.

Dankert T, K. J., Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H, van Dop PA. (2007). "A randomized clinical trial of clomifene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. ." Hum Reprod. 22: 192/197.

Dankert, T., J. A. M. Kremer, B. J. Cohlen, C. J. C. M. Hamilton, P. C. M. Pasker-de Jong, H. Straatman and P. A. van Dop (2007). "A randomized clinical trial of clomiphene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility." Human reproduction (Oxford, England) 22(3): 792-797.

Depuydt CE, D. G., Verstraete L, Vanden Broeck D, Beert JFA, Salembier G, Bosmans E, Ombelet W (2019). "Infectious human papillomavirus virions in semen reduce clinical pregnancy rates in women undergoing intrauterine insemination." Fertil Steril. 111(6): 1135-1144.

Depuydt CE, V. L., Berth M, Beert J, Bogers JP, Salembier G, Vereecken AJ, Bosmans E. (2016). "Human Papillomavirus Positivity in Women Undergoing Intrauterine Insemination Has a Negative Effect on Pregnancy Rates." Gynecol Obstet Invest 81(1): 41-46.

Diamond, M. P., R. S. Legro, C. Coutifaris, R. Alvero, R. D. Robinson, P. Casson, G. M. Christman, J. Ager, H. Huang, K. R. Hansen, V. Baker, R. Usadi, A. Seungdamrong, G. W. Bates, R. M. Rosen, D. Haisenleder, S. A. Krawetz, K. Barnhart, J. C. Trussell, D. Ohl, Y. Jin, N. Santoro, E. Eisenberg and H. Zhang (2015). "Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility." N Engl J Med 373(13): 1230-1240.

El Helw, B. (2002). Single dose letrozole versus clomiphene citrate for superovulation prior to intrauterine insemination: a prospective randomized study. M. H. El Sadek M., Fouad S., El Nomrosy K. and Abbas H., Abstracts of 18th annual meeting of ESHRE: O-209.

Erdem, M., S. Abay, A. Erdem, M. Firat Mutlu, E. Nas, I. Mutlu and M. Oktem (2015). "Recombinant FSH increases live birth rates as compared to clomiphene citrate in intrauterine insemination cycles in couples with subfertility: A prospective randomized study." European Journal of Obstetrics Gynecology and Reproductive Biology 189: 33-37.

Fatemi, H. M., E. Kolibianakis, H. Tournaye, M. Camus, A. C. Van Steirteghem and P. Devroey (2003). "Clomiphene citrate versus letrozole for ovarian stimulation: a pilot study." Reprod Biomed Online 7(5): 543-546.

Fouda, U. M. and A. M. Sayed (2011). "Extended letrozole regimen versus clomiphene citrate for superovulation in patients with unexplained infertility undergoing intrauterine insemination: A randomized controlled trial." Reproductive Biology and Endocrinology 9: no pagination.

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 Jun 21;9:84.

Galal, A. F. (2015). “Letrozole step-up protocol: The effect of a novel superovualtion induction protocol to enhance pregnancy rate in a couple with unexplained infertility undergoing intrauterine insemination.” Fertility and Sterility 104(3 SUPPL. 1): e328.

Glazener CM, Coulson C, Lambert PA, Watt EM, Hinton RA, Kelly NG, Hull MG Clomiphene treatment for women with unexplained infertility: placebo-controlled study of hormonal responses and conception rates. Gynecol Endocrinol. 1990 Jun;4(2):75-83

Goldman, M. B., K. L. Thornton, D. Ryley, M. M. Alper, J. L. Fung, M. D. Hornstein and R. H. Reindollar (2014). “A randomized clinical trial to determine optimal infertility treatment in older couples: the Forty and Over Treatment Trial (FORT-T).” Fertility and sterility 101(6): 1574-1572.

Goldman, M. B., K. L. Thornton, D. Ryley, M. M. Alper, J. L. Fung, M. D. Hornstein and R. H. Reindollar (2014). “A randomized clinical trial to determine optimal infertility treatment in older couples: The Forty and over Treatment Trial (FORT-T).” Fertility and Sterility 101(6): 1574.

Goverde, A. J., J. McDonnell, J. P. W. Vermeiden, R. Schats, F. F. H. Rutten and J. Schoemaker (2000). “Intrauterine insemination or in-vitro fertilisation in idiopathic subfertility and male subfertility: a randomised trial and cost-effectiveness analysis...including commentary by Bhattacharya S, Hall M.” Lancet 355 North American Edition(9197): 13-12.

Gregoriou, O., N. F. Vlahos, S. Konidaris, K. Papadias, D. Botsis and G. K. Creatsas (2008). “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.” Fertility and Sterility 90(3): 678-683.

Guzick, D. S., S. A. Carson, C. Coutifaris, J. W. Overstreet, P. Factor-Litvak, M. P. Steinkampf, J. A. Hill, L. Mastroianni, J. E. Buster, S. T. Nakajima, D. L. Vogel and R. E. Canfield (1999). “Efficacy of superovulation and intrauterine insemination in the treatment of infertility. National Cooperative Reproductive Medicine Network.” N Engl J Med 340(3): 177-183.

Higgins, J. P., D. Jackson, J. K. Barrett, G. Lu, A. E. Ades and I. R. White (2012). “Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies.” Res Synth Methods 3(2): 98-110.

Hutton, B., G. Salanti, D. M. Caldwell, A. Chaimani, C. H. Schmid, C. Cameron, J. P. Ioannidis, S. Straus, K. Thorlund, J. P. Jansen, C. Mulrow, F. Catala-Lopez, P. C. Gotzsche, K. Dickersin, I. Boutron, D. G. Altman and D. Moher (2015). “The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations.” Ann Intern Med 162(11): 777-784.

Ibrahim, M. I., R. A. Moustafa and A. A. Abdel-Azeem (2012). “Letrozole versus clomiphene citrate for superovulation in Egyptian women with unexplained infertility: A randomized controlled trial.” Archives of Gynecology and Obstetrics 286(6): 1581-1587.

IR, W. (2015). “Network meta analysis.” Stata journal 15: 951-985.

Kamel, A. M. (1995). Effect of induction protocols on pregnancy rate in artificial insemination by husband. Abstracts of the 11th annual meeting of ESHRE.

Karlstrom, P O Bergh, T, Lundkvist, O. A prospective randomized trial of artificial insemination versus intercourse in cycles stimulated with human menopausal gonadotropin or clomiphene citrate. Fertil Steril. 1993 Mar;59(3):554-9

Kim, D., T. Child and C. Farquhar (2015). "Intrauterine insemination: a UK survey on the adherence to NICE clinical guidelines by fertility clinics." BMJ Open 5(5): e007588.

Land, J. A. and J. L. Evers (2003). "Risks and complications in assisted reproduction techniques: Report of an ESHRE consensus meeting." Hum Reprod 18(2): 455-457.

Leanza, V., L. Coco, F. Grasso, G. Leanza, G. Zarbo and M. Palumbo (2014). "Ovulation induction with clomiphene citrate for infertile couple." Minerva Ginecol 66(3): 309-312.

Martinez, A. R., R. E. Bernardus, F. J. Voorhorst, J. P. Vermeiden and J. Schoemaker (1990). "Intrauterine insemination does and clomiphene citrate does not improve fecundity in couples with infertility due to male or idiopathic factors: a prospective, randomized, controlled study." Fertil Steril 53(5): 847-853.

Nakajima, A. K., L. L. Smith, B. Wong, J. Z. Scott, D. C. Cumming, I. V. Tataryn, M. A. Sagle, D. McAra and L. Nordstrom (1999). "A randomized trial of clomiphene citrate plus intrauterine insemination versus recombinant follicle stimulating hormone plus intrauterine insemination for the treatment of unexplained infertility." Fertility and Sterility 72(3 Suppl 1): S208.

Ombelet, W., G. Martens, P. De Sutter, J. Gerris, E. Bosmans, G. Ruyssinck, P. Defoort, G. Molenberghs and W. Gyselaers (2006). "Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study." Hum Reprod 21(4): 1025-1032.

Ozmen, B. (2005). El Sadek M., Matar H., Fouad S., El Nomrosy K. and Abbas H., Abstracts of 21st annual meeting of ESHRE.

Peeraer, K., S. Debrock, P. De Loecker, C. Tomassetti, A. Laenen, M. Welkenhuysen, L. Meeuwis, S. Pelckmans, B. W. Mol, C. Spiessens, D. De Neubourg and T. M. D'Hooghe (2015). "Low-dose human menopausal gonadotrophin versus clomiphene citrate in subfertile couples treated with intrauterine insemination: A randomized controlled trial." Human Reproduction 30(5): 1079-1088.

Peeraer, K., S. Debrock, P. De Loecker, C. Tomassetti, A. Laenen, M. Welkenhuysen, L. Meeuwis, S. Pelckmans, B. W. Mol, C. Spiessens, D. De Neubourg and T. M. D'Hooghe (2015). "Low-dose human menopausal gonadotrophin versus clomiphene citrate in subfertile couples treated with intrauterine insemination: a randomized controlled trial." Human reproduction (Oxford, England) 30(5): 1079-1088.

Robab, Davar & Asgharnia, Maryam & Tayebi, Jaber. Comparison of the success rate of letrozole and clomiphene citrate in women undergoing intrauterine insemination. Journal of Research in Medical Sciences. 2006:11.

Rumste van M.M.E., C. I. M., Veen van der F, Wely van M, Evers J.L.H. and Mol B.W.J. (2008). "The influence of the number of follicles on pregnancy rates in intrauterine insemination with ovarian stimulation: a meta-analysis." Human Rep Update 14(6).

Salanti, G., A. E. Ades and J. P. Ioannidis (2011). "Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial." J Clin Epidemiol 64(2): 163-171.

Sammour, A. (2001). Prospective randomized trial comparing the effects of letrazole (LE) and clomiphene citrate (CC) on follicular development, endometrial thickness and pregnancy rate in patients undergoing super-ovulation prior to intrauterine insemination (IUI). S. L. T. M. M. Biljan, T. Tulandi. 76: O 291.

5.

FSH or CC in IUI with ovarian stimulation for unexplained subfertility: a role for treatment selection markers?

authors:

N.A. Danhof, R. van Eekelen, S. Repping, B.W.J. Mol, F. van der Veen, M. van Wely, M.H. Mochtar, For the SUPER Study group

Reproductive BioMedicine Online. 2019;38:938-942

ABSTRACT

Research question: Can women be identified, on the basis of baseline patient characteristics, as having better chances of an ongoing pregnancy with follicle stimulation hormone (FSH) instead of clomiphene citrate as stimulation agent in intrauterine insemination for unexplained subfertility?

Design: A secondary analysis of a multicentre randomized controlled superiority trial; the SUPER study. Between July 2013 and March 2016, couples with unexplained subfertility undergoing intrauterine insemination (IUI) were allocated to an FSH or clomiphene citrate group. Female age, body mass index, duration of subfertility, primary versus secondary subfertility, antral follicle count and total motile count were assessed. For each of these factors, a logistic regression model was developed to assess if different estimated effects of FSH versus clomiphene citrate on ongoing pregnancy occurred within strata of each factor.

Results: A total of 684 couples received 2259 IUI cycles; 338 couples were allocated to FSH, of which 84 conceived leading to ongoing pregnancy and 346 couples were allocated to clomiphene citrate, of which 71 conceived leading to ongoing pregnancy. None of the treatment selection markers was associated with better ongoing pregnancy chances after IUI with FSH compared with clomiphene citrate.

Conclusion: In couples with unexplained subfertility undergoing IUI, no baseline treatment selection markers could be identified to determine whether ovaries should be stimulated with FSH or clomiphene citrate.

Study funding/competing interest(s): The initial trial was funded by the Netherlands Organization for Health Research and Development (ZonMw). (Health Care Efficiency Research; project number: 80-83600-98-10192). The Eudract number for this trial was 2013-001034-18.

INTRODUCTION

Intrauterine insemination (IUI) with ovarian stimulation is recommended as first-line treatment in couples diagnosed with unexplained infertility and men with a total motile sperm count of over 10 million and a prognosis of spontaneous pregnancy less than 30% within a year (Practice committee ASRM, 2006, Kim et al., 2015, Calhaz-Jorge et al., 2017, Cohlen et al., 2018).

Ovarian stimulation in IUI can be carried out with follicle stimulation hormone (FSH), clomiphene citrate or letrozole. Letrozole has not been registered for this indication, which is why FSH and clomiphene citrate are still standard medications for ovarian stimulation in many countries. Several studies have compared the effectiveness of FSH with clomiphene citrate (Karlstrom et al., 1993, Balasch et al., 1994, Kamel, 1995, Karlstrom et al., 1998, Ecochard et al., 2000, Dankert et al., 2007, Nakajima et al., 1999. The studies that used high doses of FSH generally found higher pregnancy rates after FSH compared with clomiphene citrate. Consequently, a Cochrane review suggested that pregnancy rates were higher with FSH, although data on multiple pregnancies were too limited to draw any conclusions (Cantineau and Cohlen, 2007). A subsequent, large multicentre trial provided insight into multiple pregnancy rates, which ranged from 10% for FSH and 3% for clomiphene citrate (Diamond et al., 2015). This high multiple pregnancy is not acceptable in view of the high risks of maternal and neonatal complications.

We recently completed a multicentre randomized controlled trial investigating the effectiveness of FSH compared with clomiphene citrate in couples with unexplained subfertility undergoing IUI in a protocol with strict cancellation criteria (Danhof et al., 2018). We found multiple pregnancy rates of 1% (5/369) after FSH and 2% (8/369) after clomiphene citrate (RR 0.63, 95% CI 0.21 to 1.89), without compromising on cumulative ongoing pregnancy rates, which were 31% (113/369) after FSH and 26% (97/369) after clomiphene citrate (RR 1.16, 95% CI 0.93 to 1.47) (Danhof et al., 2018).

Although these data clearly indicate that risk of the multiple pregnancy (1–2% per woman) is low with this protocol, it also clearly shows that most couples (69–74%) did not conceive within 6 months of treatment. This raises the question of whether ongoing pregnancy rates would have been higher after FSH or clomiphene citrate for some subgroups of couples.

The aim of this study was, therefore, to assess whether women can be identified, on the basis of baseline patient characteristics, as having better chances of an ongoing pregnancy with FSH instead of clomiphene citrate as stimulation agent in IUI for unexplained subfertility.

MATERIALS AND METHODS

Study design

A secondary analysis of the SUPER study was conducted to assess whether baseline patient characteristics were associated with better ongoing pregnancy chances after IUI with FSH compared with clomiphene citrate. The SUPER study was a multicentre randomized controlled trial conducted by the Dutch Consortium for Healthcare Evaluation in Obstetrics and Gynaecology (www.zorgevaluatienederland.nl). It compared FSH with clomiphene citrate in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria (Danhof et al., 2018). Ethical approval for this study was obtained from the Medical Ethical Committee of the Academic Medical Centre and from the Dutch Central Committee on Research involving Human Subjects on 16 April 2013 (CCMO NL 43131-018-13). The Board of Directors of each of the participating centres approved local execution of the study. In this study, we briefly discuss the trial essentials, as more in-depth information has already been published (Danhof et al., 2018).

Couples diagnosed with unexplained subfertility, defined as failure to conceive after 1 year of regular unprotected intercourse, a prewash total motile sperm count of at least 3 million and at least one side tube patency proven by sonohysterography, hysterosalpingography or by diagnostic laparoscopy, were scheduled for a maximum of four cycles of IUI with ovarian stimulation comparing 75 IU FSH with 100 mg clomiphene citrate. In the first treatment cycle, all women were seen at a baseline visit for a transvaginal ultrasound examination on the third, fourth or fifth day of the menstrual cycle. Women were not allowed to start the treatment cycle if one or more ovarian cysts of over 20 mm were seen. In the FSH treatment arm, women started with daily subcutaneous injections of 75 IU FSH on days 3, 4 or 5 of the menstrual cycle and continued these injections until the day of ovulation triggering. In the clomiphene citrate treatment arm, women started with 100 mg clomiphene citrate on days 3, 4 or 5 of the menstrual cycle. The tablets were administered orally and were stopped after 5 days of daily intake. In both interventions, follicular development was monitored by transvaginal ultrasound. The total number of follicles and their diameters were registered. Ovulation was triggered with 5000 IU HCG or with 250 µg recombinant HCG if there was at least one dominant follicle with a mean diameter of 16–18 mm and a maximum of three follicles measuring 14 mm or wider. Strict cancellation criteria were applied to reduce the number of multiple pregnancies. Intrauterine insemination was cancelled on cycle day if more than three follicles with a diameter of 14 mm or wider or five follicles with a diameter of 12 mm or wider were seen at transvaginal ultrasound.

Statistical analysis

Data were analysed on cycle level. Couples who started IUI using their allocated stimulation agent were included, thus eliminating couples who conceived naturally before starting ovarian stimulation. Female age, body mass index (BMI), duration of subfertility, whether a couple was primary or secondary subfertile, total motile count

(TMC) and antral follicle count (AFC) were assessed as potential treatment selection markers. For each of these markers, a logistic regression model was developed with ongoing pregnancy after IUI as the outcome and the marker, and the allocated drug and a factor-by-drug interaction term as predictors. The interaction indicates whether the effect of FSH compared with clomiphene citrate is dependent on that factor. In the case of interaction, the probability of an ongoing pregnancy over the range of the identified treatment selection factor was graphically expressed. Couples receiving multiple cycles were accounted for by adjusting precision measures using a robust variance estimator (Berhane and Weissfeld, 2003). Missing data were accounted for by multiple imputation. All numerical results are based on pooled estimates over 10 imputation sets using Rubin's Rules (Rubin, 2004). $P < 0.05$ indicated statistical significance. SPSS version 24 and R version 3.3.2 was used for statistical analyses.

RESULTS

Study group

Between July 2013 and March 2016, 369 women were randomly allocated to ovarian stimulation with FSH and 369 women to ovarian stimulation with clomiphene citrate. After exclusion of natural conceptions and couples switching stimulation agent before the first cycle, 684 couples remained who received 2259 IUI cycles in total, of whom 338 couples were allocated to FSH and 346 to clomiphene citrate. Baseline characteristics of couples randomized to FSH were well balanced with those randomized to clomiphene citrate (Table 1). The rates of ongoing pregnancy per cycle are presented in Table 2. A total of 84 women conceived leading to ongoing pregnancy, i.e. a rate per cycle of 7.6% in the FSH treatment arm and 71 women conceived leading to ongoing pregnancy, i.e. a rate per cycle of 6.2% in the clomiphene citrate treatment arm.

Table 1. Baseline characteristics

Characteristics	FSH (*n*=330)	Clomiphene citrate (*n*=346)
Mean female age (years)	33·3 ± 4.0	33·6 ± 4·0
Primary subfertility	248 (73)	252 (73)
Median duration of subfertility (months)	25 (20 - 33)	24 (20 - 32)
Median body mass index, kg/m²	23·3 (21·1 - 26·0)	23·1 (21·0 - 25·4)
Median total motile count, x10^6	48 (22 – 98)	59 (26 - 116)
Antral follicle count	13 (9-20)	13 (9-19)

* Data are n (%), mean (SD) or median (25th – 75th percentiles)

Table 2. Ongoing pregnancies per cycle

Cycle number	FSH (%, 95%CI)	Clomiphene citrate (%, 95%CI)
1	28/338 (8.3, 5.7-11.9)	26/346 (7.5, 5.1-10.9)
2	22/286 (7.7, 5.0-11.6)	18/302 (6.0, 3.7-9.4)
3	13/243 (5.3, 3.0-9.2)	17/255 (6.7, 4.1-10.6)
4	17/195 (8.7%, 5.3-13.8)	8/209 (3.8, 1.8-7.7)
5 to 7	4/44 (9.1, 3.0-22.6)	2/41 (4.9, 0.8-17.8)

The mean number of follicles measuring 14 mm or wider on day of ovulation triggering was 1.8 (SD 1.43) in the FSH treatment arm and the mean number of follicles measuring 14 mm or wider on day of ovulation triggering was 1.9 (SD 1.11) in the clomiphene citrate treatment arm, indicating no statistical significance. No difference was observed in the cancellation rate owing to the development of more than three dominant follicles between ovarian stimulation with FSH and ovarian stimulation with clomiphene citrate (FSH: n = 115; clomiphene citrate: n = 101; RR 1.17; 95% CI 0.93 to 1.45).

Treatment selection markers

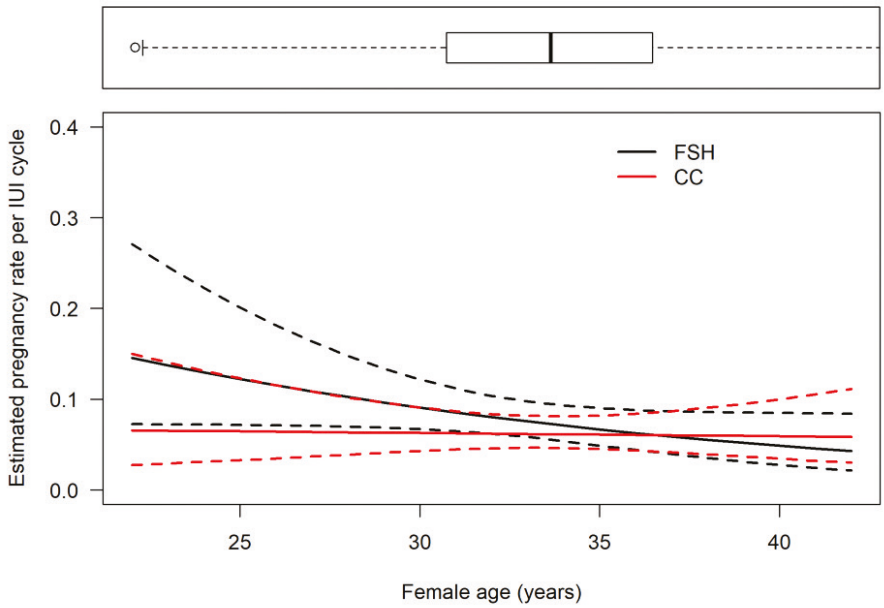
The associations between potential treatment selection markers and the chances of an ongoing pregnancy are presented in Table 3. For BMI, duration of subfertility, primary or secondary subfertility, TMC, AFC and age, the confidence intervals of FSH and clomiphene citrate overlapped, indicating no statistically significant association. None of these markers were associated with higher chances of an ongoing pregnancy with ovarian stimulation with FSH compared with ovarian stimulation with clomiphene citrate or vice versa (Table 3). The non-statistically significant interaction between female age and the probability of an ongoing pregnancy after FSH or clomiphene citrate is shown in Figure 1. The probability of an ongoing pregnancy after IUI with FSH was 14.5% per cycle for a female age of 22 years and decreased to 4.3% per cycle for a female age of 42 years. The probability of an ongoing pregnancy was 6.5% per cycle after IUI with clomiphene citrate, regardless of female age.

Table 3. The association between potential treatment selection markers and ongoing pregnancy

Potential treatment selection markers	After FSH OR, 95%CI	After clomiphene citrate OR, 95%CI
Female age, years	0.94 (0.87-1.02)	0.99 (0.94-1.05)
BMI, one unit	1.04 (0.96-1.12)	1.01 (0.95-1.07)
Duration of subfertility, years	1.14 (0.81-1.53)	0.96 (0.72-1.28)
Primary versus secondary subfertile	1.09 (0.49-2.38)	1.33 (0.75-2.37)
TMC, per 10 x 10^6	1.02 (0.98-1.06)	0.99 (0.96-1.02)
Antral follicle count	1.00 (0.99-1.01)	1.00 (0.99-1.00)

All P-values for interaction were not statistically significant.

Figure 1. Female age and the probability of an ongoing pregnancy for either FSH or clomiphene citrate (CC). The age distribution is shown as a box-plot above the graph. IUI, intrauterine in-**semination. Female age**



DISCUSSION

In women with unexplained subfertility, undergoing IUI with ovarian stimulation, the boundaries of the 95% confidence interval of the associations with each intervention overlapped for all investigated treatment selection markers, indicating no significant difference between FSH and clomiphene citrate. Hence, we could not identify any couples who were about to start IUI and ovarian stimulation that would have better chances of an ongoing pregnancy with ovarian stimulation with FSH instead of clomiphene citrate.

A strength of this study is that the analyses were based on the results of a well-powered multicentre randomized controlled trial in which the baseline characteristics were well balanced between the treatment arms. This provides the opportunity to investigate interactions between baseline characteristics and ongoing pregnancy without the uncertainty of selection bias.

A potential limitation of this study is that secondary analyses should be interpreted with caution as they are prone to false positive findings. In addition, our sample size was rather limited for a secondary analysis, which makes it impossible to detect smaller

associations between potential treatment selection markers and ongoing pregnancy. On the other hand, the clinical relevance of any smaller associations might be questioned.

The focus from one-size-fits all strategies has shifted somewhat into personalized treatment based on specific patient characteristics (Schork, 2015). In reproductive medicine, patient characteristics that predict the chances of natural conception in the next year have been identified and are widely used, but these patient characteristics have not yet been translated into specific treatment recommendations (van der Steeg et al., 2007, van Eekelen et al., 2017). We chose to investigate baseline characteristics that are important factors in the prediction of conception. The lack of any association between these potential treatment selection markers at baseline and ongoing pregnancy rates may point to the underlying mechanism of multifollicular growth. The generation of multi-follicular growth is considered to be responsible for the improvement in ongoing pregnancy rates after IUI (van Rumste et al., 2008, Cohlen et al., 2018). A strong correlation exists between the number of dominant follicles developed and ongoing pregnancy rates (van Rumste et al., 2006, van Rumste et al., 2008, Danhof et al., 2018). In the SUPER study, numbers of dominant follicles were equal in both treatment arms (Danhof et al., 2018). On the basis of our data, similar numbers of dominant follicles lead to equal pregnancy rates regardless of differences in the way FSH and clomiphene citrate exert their effect.

Our finding of no treatment selection markers to decide which stimulation agent should be used in couples with unexplained subfertility undergoing IUI is in line with other studies on personalized treatment in couples with unexplained subfertility. In 2017, a study reported that no treatment selection markers were found that could identify couples that would have better chances of a live birth with IVF compared with IUI with ovarian stimulation (Tjon-Kon-Fat et al., 2015). For that reason, the investigators concluded that IUI with ovarian stimulation should remain the first-line treatment in all couples with unexplained subfertility. The issue of individualized FSH dosing in women who are about to start with IVF and intracytoplasmic sperm injection was reported in two trials that found that this does not improve live birth rates (Oudshoorn et al., 2017, van Tilborg et al., 2017).

In conclusion, we could not identify treatment selection markers at baseline in couples with unexplained subfertility to decide on ovarian stimulation with FSH instead of clomiphene citrate in IUI. We suggest that a decision to opt for a stimulation agent is made in consultation with the couples based on effectiveness, costs and their preference.

REFERENCES

Balasch J, Ballesca JL, Pimentel C, Creus M, Fabregues F, Vanrell JA. Late low-dose pure follicle stimulating hormone for ovarian stimulation in intra-uterine insemination cycles. Human Repdroduction 1994;9(10):1863–6

Berhane K, Weissfeld LA. Inference in spline-based models for multiple time-to-event data, with applications to a breast cancer prevention trial. Biometrics. 2003 Dec;59(4):859-68.

Calhaz-Jorge C, De Geyter C, Kupka M, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C and Goossens V. European IVF-monitoring Consortium (EIM); European Society of Human Reproduction and Embryology (ESHRE), Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod. 2017. 10:1957-1973

Cantineau AEP and Cohlen BJ. Ovarian stimulation protocols (anti-oestrogens, gonadotrophins with and without GnRH agonists/antagonists) for intrauterine insemination (IUI) in women with subfertility. Cochrane Database Syst Rev. 2007

Cohlen B, Bijkerk A, Van der Poel S, Ombelet W. IUI: review and systematic assessment of the evidence that supports global recommendations. Hum Reprod Update. 2018 May 1;24(3):300-319.

Danhof NA, van Wely M, Repping S, Koks C, Verhoeve HR, de Bruin JP, Verberg MFG et al. Follicle stimulating hormone (FSH) versus Clomiphene Citrate (CC) in intrauterine insemination for unexplained subfertility: A randomized controlled trial. Hum Reprod. 2018; dey268

Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H and van Dop PA. A randomized clinical trial of clomifene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. Hum Reprod. 2007;3:792-7

Diamond MP, Legro RS, Coutifaris C, Alvero R, Robinson RD, Casson P, Christman GM, Ager J, Huang H, Hansen KR, Baker V, Usadi R, Seungdamrong A, Bates GW, Rosen RM, Haisenleder D, Krawetz SA, Barnhart K, Trussell JC, Ohl D, Jin Y, Santoro N, Eisenberg E, Zhang H; NICHD Reproductive Medicine Network. Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility. N Engl J Med. 2015;13:1230-40

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. Fertility and Sterility 2000;73(1):90–3.

Kamel MA. Effect of induction protocols on pregnancy rate in artificial insemination by husband. Human Reproduction. 1995; Vol. 10 Abstractbook:116

Karlstrom PO, Bergh T, Lundkvist O. A prospective randomized trial of artificial insemination versus intercourse in cycles with human menopausal gonadotropin or clomiphene citrate. Fertility and Sterility 1993;59(3):554–9.

Karlstrom PO, Berkurezion M, Bergh T, Lundkvist O. An extended prospective randomized trial of artificial insemination versus intercourse in cycles stimulated with human menopausal gonadotrophins (hMG) or clomiphene citrate (CC). Fertility and Sterility 1998;70(3):Abstract book Suppl 2, S420.

Kim D, Child T, Farquhar C. Intrauterine insemination: a UK survey on the adherence to NICE clinical guidelines by fertility clinics. BMJ Open. 2015; 5(5):e007588

Nakajima AK, Smith LL, Wong B, Scott JZ, Cumming DC, Tataryn IV, Sagle MA, McAra D, Nordstrom L. A randomized trial of clomiphene citrate plus intrauterine insemination versus recombinant follicle stimulating hormone plus intrauterine insemination for the treatment of unexplained infertility. Fertility and Sterility. 1999; Vol. 72 Suppl 1 Abstract book 3:S208.

Oudshoorn SC, van Tilborg TC, Eijkemans MJC, Oosterhuis GJE, Friederich J, van Hooff MHA, van Santbrink EJP, Brinkhuis EA, Smeenk JMJ, Kwee J, de Koning CH, Groen H, Lambalk CB, Mol BWJ, Broekmans FJM, Torrance HL; OPTIMIST study group. Individualized versus standard FSH dosing in women starting IVF/ICSI: an RCT. Part 2: The predicted hyper responder. Hum Reprod. 2017;32(12):2506-2514

The Practice Committee of the American Society for Reproductive Medicine Effectiveness and treatment for unexplained infertility Fertil Steril 2006; 86(4):S111-S114

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

Schork NJ. Personalized medicine: Time for one-person trials. Nature. 2015 Apr 30;520(7549):609-11

Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A, Hompes P, Broekmans FJ, Verhoeve HR, de Bruin JP, van Golde R, Repping S, Cohlen BJ, Lambers MD, van Bommel PF, Slappendel E, Perquin D, Smeenk J, Pelinck MJ, Gianotten J, Hoozemans DA, Maas JW, Groen H, Eijkemans MJ, van der Veen F, Mol BW, van Wely M. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015;10:2331-9.

van Rumste MM, den Hartog JE, Dumoulin JC, Evers JL, Land JA. Is controlled ovarian stimulation in intrauterine insemination an acceptable therapy in couples with unexplained non-conception in the perspective of multiple pregnancies? Hum Reprod. 2006 Mar;21(3):701-4.

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 Nov-Dec;14(6):563-70

van der Steeg JW1, Steures P, Eijkemans MJ, Habbema JD, Hompes PG, Broekmans FJ, van Dessel HJ, Bossuyt PM, van der Veen F, Mol BW; CECERM study group (Collaborative Effort for Clinical Evaluation in Reproductive Medicine). Pregnancy is predictable: a large-scale prospective external validation of the prediction of spontaneous pregnancy in subfertile couples. Hum Reprod. 2007 Feb;22(2):536-42. Epub 2006 Sep 22.

van Eekelen R, Scholten I, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Hompes P, van Wely M, van der Veen F, Mol BW, Eijkemans MJ, Te Velde ER, van Geloven N. Natural conception: repeated predictions over time. Hum Reprod. 2017 Feb;32(2):346-353.

6.

Endometrial thickness as a biomarker for ongoing pregnancy in intrauterine insemination for unexplained subfertility: a secondary analysis

authors:

N.A. Danhof, R. van Eekelen, S. Repping, B.W.J. Mol, F. van der Veen, M. van Wely, M.H. Mochtar, For the SUPER Study group

Accepted for publication in Human Reproduction Open

ABSTRACT

Study question: What is, in couples with unexplained subfertility undergoing IUI, the impact of gonadotrophins compared to clomiphene citrate (CC) on endometrial thickness (EMT) in relation to ongoing pregnancy?

Summary answer: In women with unexplained subfertility undergoing IUI with ovarian stimulation, gonadotrophins lead to a thicker endometrium compared to CC, but this does not affect ongoing pregnancy rates.

What is known already: A systematic review and meta-analysis among couples with unexplained subfertility undergoing IUI with ovarian stimulation showed that women who conceived had, on average, a thicker endometrium than women who did not conceive, but this evidence is not robust due to a high level of heterogeneity. There was insufficient data to draw any conclusions on EMT and the effect on pregnancy outcomes.

Study design, size, duration: We performed a secondary analysis of a multicentre randomised controlled superiority trial in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria. In total, 738 couples recruited between July 2013 and March 2016 were allocated to ovarian stimulation with gonadotrophins (n=369) or with CC (n=369) for a maximum of four IUI cycles. According to local protocol, recombinant FSH, urinary FSH or hMG was used. Natural conceptions and cancelled cycles were removed from this secondary analysis, as they do not provide any information on pregnancy in relation to stimulation after IUI. Ongoing pregnancy was defined as a positive heartbeat at or beyond 12 weeks of gestation.

Participants/materials, settings, methods: We first determined the difference in EMT between women randomised to gonadotrophins (75 IU) and CC (100mg) over all cycles using a linear mixed model. We then investigated the association between EMT and ongoing pregnancy after IUI using a logistic regression model, adjusted for the allocated drug, number of dominant follicles, female age, BMI, duration of subfertility, primary or secondary subfertility, referral status, smoking status, cycle number and total motile sperm count. To conclude, we investigated the association between EMT and ongoing pregnancy by logistic regression separately in women allocated to gonadotrophins and in women allocated to CC.

Main results and the role of chance: A total of 666 couples underwent 1968 IUI cycles. Of these, 330 couples were allocated to gonadotrophins, of which 85 conceived leading to ongoing pregnancy (rate per cycle 8.9%) and 336 couples were allocated to CC, of which 71 conceived leading to ongoing pregnancy (rate per cycle 7.0%) (relative risk (RR) 1.22, 95% CI 0.92 to 1.61). The mean EMT was 8.9 mm (SD 2.1) in women treated with gonadotrophins and 7.5 mm (SD 2.1) in women treated with CC (adjusted mean difference 1.4 mm; 95% CI: 1.1 to 1.7). The overall mean EMT was 8.4 mm (SD 2.2) in women that conceived leading to ongoing pregnancy and 8.2 mm (SD 2.2) in women

that did not conceive (adjusted odds ratio (OR): 1.03 per 1 mm increase, 95% CI 0.95 to 1.12). There was no association between EMT and ongoing pregnancy in women treated with gonadotrophins or CC (OR:1.01 per 1 mm increase, 95% CI 0.90 to 1.13, and 1.10 per 1 mm increase, 95% CI 0.99 to 1.23, respectively).

Limitation, reason for caution: Since this is a secondary analysis, the data should be interpreted prudently as secondary analyses are prone to false positive findings or could be underpowered to show associations that the study is not primarily set up for.

Wider implications of the findings: In women with unexplained subfertility and treated with IUI, gonadotrophins lead to a significantly thicker endometrium compared to CC, but there was no evidence of a consistent association between EMT in women treated with gonadotrophins or CC and the ongoing pregnancy rate. A relatively thin endometrium after CC is therefore not a valid reason to prefer gonadotrophins as the stimulation agent in IUI for unexplained subfertility.

Study funding/competing interest(s): The initial trial was funded by the Netherlands Organization for Health Research and Development (ZonMw) (Health Care Efficiency Research; project number: 80-83600-98-10192). The Eudract number for this trial was 2013-001034-18.

Prof dr BWJ Mol is supported by a NHMRC Practitioner Fellowship (GNT1082548). BWM reports consultancy for Merck, ObsEva and Guerbet. The other authors declare no conflicts of interest.

Trial registration: NTR 4057

What does this mean for patients: Intrauterine insemination (IUI) combined with stimulation of the woman's ovaries is often used for couples with unexplained subfertility. Ovarian stimulation can be done with gonadotrophins injections (e.g. LH, FSH, which normally stimulate egg development) or tablets of clomiphene citrate (a drug which causes LH/FSH release).

Studies have shown that IUI with ovarian stimulation with gonadotrophins leads to a 5% increase in ongoing pregnancy rates compared to IUI with clomiphene citrate. We proposed that this difference in ongoing pregnancy rate might be explained by a difference in thickness of the endometrium: the endometrium is the inner lining of the uterus, which thickens each month in preparation for a possible pregnancy.

In this study, we performed a secondary analysis of data from patients who underwent IUI and found that injecting gonadotrophins led to a thicker endometrium compared to tablets of clomiphene citrate. However, this difference in endometrial thickness did not affect the ongoing pregnancy rate. Therefore a relatively thin endometrium after clomiphene citrate is not a reason to prefer gonadotrophins as the stimulation agent in IUI for unexplained subfertility.

INTRODUCTION

In numerous countries, a first line treatment for couples diagnosed with unexplained subfertility is IUI with ovarian stimulation (Calhaz-Jorge et al., 2017; The Practice Committee of the American Society for Reproductive Medicine., 2006). In Europe, 175,000 IUI cycles are performed each year (Calhaz-Jorge et al. 2017).

Ovarian stimulation in IUI can be performed with s.c. injections of gonadotrophins or with clomiphene citrate (CC) tablets taken orally. Increased serum gonadotrophin levels during the follicular phase induces growth of more than one dominant follicle, while CC occupies the estrogen receptors, blocks the negative feedback of 17β-estradiol in the hypothalamic pituitary axis and thereby indirectly increases serum gonadotrophin levels (Schipper et al 1998; Wu 1977; Kettel et al 1993). Although the concept of ovarian stimulation is that it is the development of multiple follicles that increases pregnancy rates, ovarian stimulation also affects endometrial thickness (EMT).

In a recent multicentre randomized controlled trial comparing gonadotrophins to CC in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria, the ongoing pregnancy rate was 31% in women allocated to gonadotrophins and 26% in women allocated to CC (relative risk (RR) 1.16, 95% CI 0.93 to 1.47), while there was no evidence of a statistically significant difference in the mean ±SD number of follicles ≥14 mm at the day of ovulation triggering (gonadotrophins 1.8±1.43, CC 1.9±1.11, p=0.52) (Danhof et al., 2018). The question then arises of whether this difference in ongoing pregnancy rates can be explained by a difference in EMT.

To investigate the association between EMT and pregnancy in women with unexplained subfertility undergoing IUI, a systematic review and meta-analysis pooled the data of two randomised controlled trials (RCTs) and five cohort studies and found a thinner endometrium after CC as compared to gonadotrophin stimulation, although the difference was not statistically significant (mean difference: 0.51 mm, 95% CI: −0.05 to 1.07; I2 = 74%) (Weiss et al 2017). When comparing various drugs used for ovarian stimulation, the evidence was insufficient to draw any conclusions on the impact of the type of drug on EMT and on pregnancy outcomes (Weiss et al 2017). Based on the existing evidence, it is unclear whether a difference in EMT following different ovarian stimulation agents in IUI can lead to a difference in pregnancy outcomes.

We therefore performed a secondary analysis of our multicentre RCT, to explore the impact of gonadotrophins compared to CC on EMT and its possible impact on ongoing pregnancy.

METHODS

Study design

We conducted a secondary analysis of the SUPER study (Danhof et al. 2018). Here, we briefly discuss the trial essentials as details have been described elsewhere (Danhof et al. 2018). The Medical Ethical Committee of the Academic Medical Centre and the Dutch Central Committee on Research involving Human Subjects approved this study (CCMO NL 43131-018-13) and the board of directors of each participating site approved local execution (NTR4057). Couples diagnosed with unexplained subfertility were scheduled for a maximum of four IUI cycles with ovarian stimulation comparing 75 IU gonadotrophins to 100 milligrams CC within a time horizon of 6 months. According to local protocol, either recombinant FSH, urinary FSH or hMG was used. Women could thus receive multiple treatment cycles. In both interventions, we monitored follicular development and EMT by transvaginal ultrasound. We cancelled insemination if more than three follicles with a diameter of ≥14 mm or five follicles with a diameter of ≥12 mm were seen at transvaginal ultrasound, regardless of the EMT. EMT was not a criterion to cancel the cycle or switch medication. Ongoing pregnancy was defined as a positive heartbeat at or beyond 12 weeks of gestation.

Statistical analysis

Data were analysed on cycle level. Natural conceptions and cancelled cycles were removed from analysis, as they are not informative on pregnancy rates after IUI in relation to EMT. First, we determined the difference in EMT between women randomised to gonadotrophins and CC over all cycles with a linear mixed model and EMT as the outcome. We handled the allocated drug as a fixed covariate, and used random intercepts and random slopes for cycle number, taking into account that women could receive multiple treatment cycles. Second, we investigated the estimated trend in EMT over subsequent cycles for individual women. Third, we determined the difference in EMT between women that had an ongoing pregnancy and women that did not. Fourth, we investigated the association between EMT and ongoing pregnancy after IUI using a logistic regression model. In this model, we adjusted for the stimulation agent, female age, duration of subfertility, primary or secondary subfertility, referral status, BMI, total motile sperm count (TMSC), smoking status, the growth of two follicles versus one follicle, or three follicles versus one follicle, and for failed IUI cycles. Finally, we investigated the association between EMT and ongoing pregnancy rate by logistic regression in women allocated to gonadotrophins and in women allocated to CC.

Missing data

We decided on multiple imputation since missing data on EMT occurred in 68 out of 1968 cycles (3.5%) and multiple imputation is generally advised in these settings.

Numerical results are based on pooled estimates over 10 imputation sets using Rubin's Rules (Rubin, 2004).

We used SPSS Version 22.0 (IBM Software United States) and R version 3.3.2 (R Core Team (2016).

RESULTS

Study group

Between July 2013 and March 2016, we randomly allocated 369 women to ovarian stimulation with gonadotrophins and 369 women to ovarian stimulation with CC. After exclusion of natural conceptions and cancelled cycles, 666 couples remained who underwent 1968 IUI cycles in total.

A total of 330 couples were allocated to gonadotrophins and 336 couples to CC. The baseline characteristics were well balanced between couples that were allocated to gonadotrophins or CC (Table I). In the gonadotrophin treatment arm, 85 women conceived leading to ongoing pregnancy (ongoing pregnancy rate per cycle 8.9%), while 71 women conceived leading to ongoing pregnancy (ongoing pregnancy rate per cycle 7.0%) in the CC treatment arm. In the gonadotrophin treatment arm the median [interquartile range (IQR) number of follicles was 1.67 (1) and in the CC treatment arm this was 1.75 (1).

Table 1. Baseline characteristics of the participating couples with unexplained subfertility undergoing IUI in the present study.

Characteristics	Gonadotrophins* (*n*=330)	CC (*n*=336)
Mean female age (years)	33·3 ± 4.0	33·7 ± 4·0
Primary subfertility	243 (74%)	245 (73%)
Median duration of subfertility (months)	24 (20 - 33)	24 (19 - 31)
Current smoking status	55 (17%)	49 (15%)
Median BMI in kg/m²**	23·3 (21 - 26)	23·1 (21 - 25)
Median total motile count (x10^6)	48 (22 - 98)	59 (27 - 113)

Data are n (%), mean (SD) or median (25th-75th percentiles)
* recombinant FSH, urinary FSH or hMG was used
**BMI is the weight in kilograms divided by the square of height in meter
CC=clomiphene citrate

EMT

The mean EMT was 8.9 mm (SD 2.1) in women treated with gonadotrophins and 7.5 mm (SD 2.1) in women treated with CC (adjusted mean difference 1.4 mm; 95%CI: 1.1-1.7). The difference in EMT over subsequent cycles was on average 0.02 mm per additional

cycle (Fig. 1) and ranged from -0.6 to 0.6. This suggests that there are no clear trends of EMT over multiple cycles for the same woman.

In ovarian stimulation with gonadotrophins there was no statistically significant difference between EMT and ongoing pregnancy (odds ratio (OR): 1.01 per 1 mm increase, 95%CI 0.90-1.13). In ovarian stimulation with CC there was also no statistically significant difference between EMT and ongoing pregnancy (OR: 1.10 per 1 mm increase, 95%CI 0.99-1.23).

Adjusting for known predictors of pregnancy, such as the stimulation agent, female age, duration of subfertility, primary or secondary subfertility, referring status, BMI, TMSC, smoking status, the growth of two follicles versus one follicle, or three follicles versus one follicle, and for failed IUI cycles, did not change results (Table II).

Table 2. Results for the adjusted logistic regression model.

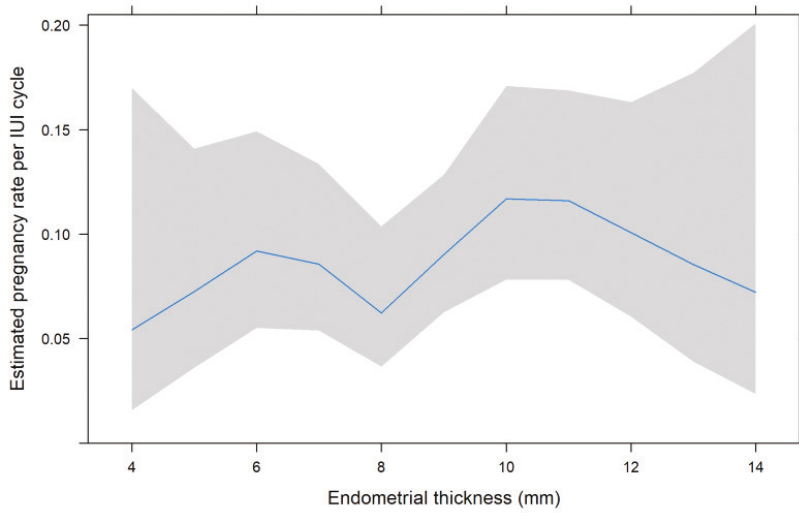
	OR for ongoing pregnancy after IUI	95% CI
EMT, per mm	1.04	0.95-1.12
Female age, per year	0.96	0.92-1.00
Duration of subfertility, per year	1.02	0.87-1.21
Primary versus secondary subfertility	1.40	0.93-2.09
Referred by Ob/Gyn versus referred by GP	0.78	0.45-1.35
BMI, per unit	1.04	1.00-1.08
Total motile sperm count, per $10x10^6$	1.00	0.98-1.02
Gonadotrophins versus CC	1.27	0.89-1.82
Smoking, yes versus no	1.01	0.64-1.60
Two follicles versus one follicle	1.59	1.10-2.30
Three follicles versus one follicle	1.99	1.24-3.21
Per failed IUI cycle	0.89	0.77-1.03

CC=clomiphene citrate, EMT=endometrial thickness, OR=odds ratio Ob/Gyn: obstetrics/gynaecology

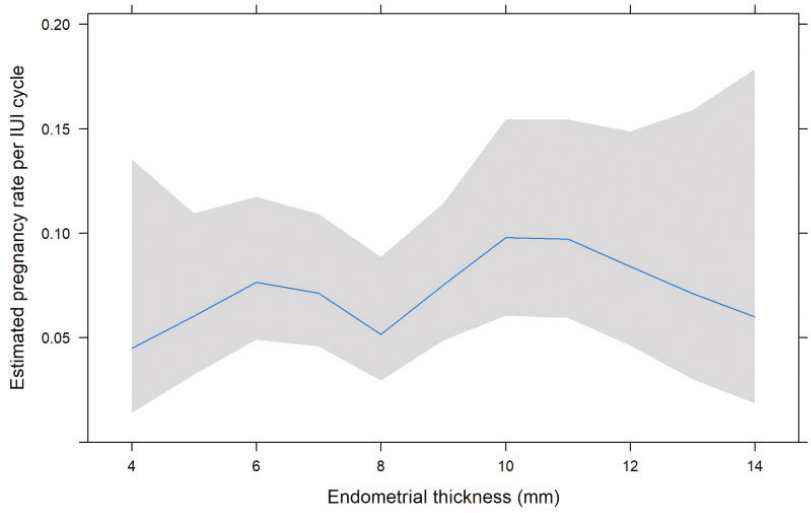
The non-linear association between EMT and the estimated chance of ongoing pregnancy after IUI with, respectively, gonadotrophins and CC is shown (Fig. 2a and b). There was no statistically significant association between EMT and ongoing pregnancy in cycles with ovarian stimulation with gonadotrophins, or cycles with ovarian stimulation with CC (Fig. 2).

Figure 2. The non-linear association between EMT and the estimated chance of ongoing pregnancy

A. The non-linear association between EMT and the estimated chance of ongoing pregnancy after IUI with gonadotrophins



B. The non-linear association between EMT and the estimated chance of ongoing pregnancy after IUI with CC



EMT=endometrial thickness

DISCUSSION

This study demonstrates that in women with unexplained subfertility undergoing IUI with ovarian stimulation EMT and ongoing pregnancy were not associated, regardless of whether they were treated with gonadotrophins or CC (OR:1.01 per 1 mm increase, 95%CI 0.90-1.13, and OR 1.10 per 1 mm increase, 95%CI 0.99-1.23, respectively).

A strength of this study is that this analysis is based on the results of the first multicentre RCT randomizing between gonadotrophins and CC in IUI for unexplained subfertility in which EMT is measured. This is the largest study so far that investigated the association between EMT and ongoing pregnancy in women undergoing IUI for unexplained subfertility and that compared two stimulation regimens. A potential limitation of this study is that this is a secondary analysis and should thus be interpreted prudently, as secondary analyses are prone to false positive findings or could be underpowered to show associations that the study is not primarily set up for. Also, we cannot exclude that a difference of 1.4 mm in EMT is due to interobserver variability or due to differences in equipment. Differences in EMT measurements of 1.5 mm have been reported between experienced and inexperienced transvaginal sonography examiners (Karlsson et al 1994). The impact of this is most likely to be limited, since our study is based on a multicentre RCT with women from 24 clinics, thereby reflecting daily clinical practice.

We found a thinner endometrium in women who had ovarian stimulation with CC, which can be explained by the anti-estrogenic effect of CC, since endometrium proliferates under the influence of estrogen. In histological studies on the effect of CC on the development of the endometrium, CC had a deleterious effect on the maturity of the endometrium (Yeko et al 1992; Massai et al 1993; Unfer et al 2001). The relation between these histological changes and the EMT measured by ultrasound has not been clarified yet (Zaidi et al 1995; Unfer et al 2000). Nevertheless, we feel histology has a very limited impact, since the average EMT was very similar for women who conceived and women who did not conceive and there was no association between EMT and ongoing pregnancy when looking at women allocated to either gonadotrophins or CC. Since the sample size was too small to draw any firm conclusions on EMTs of less than 4 mm or more than 12 mm, uncertainty remains in these cases.

Our data regarding the difference in EMT as a result of ovarian stimulation with either gonadotrophins or CC and the lack of any association with ongoing pregnancy are in agreement with a recently performed systematic review and meta-analysis (CC versus gonadotrophins; two studies, mean difference: -0.33, 95% CI: -0.64 to -0.01) (Weiss et al 2017). Our study also confirms the results of another systematic review on EMT and pregnancy rates performed in women with World Health Organization group II anovulation, with respect to a thinner endometrium after CC compared to gonadotrophins (one study, weighted mean differences [WMD] -0.08, 95% CI -0.89 to 0.73) and no evidence of a significant difference in the association between EMT and pregnancy (one study, RR 0.71, 95% CI 0.23 to 2.15) or live birth (one study, RR 0.85,

95% CI 0.26 to 2.73) (Gadalla et al 2018). Our findings are in contrast with a systematic review and meta-analysis on EMT and pregnancy rates after IVF, which found that the chance of a clinical pregnancy with an EMT ≤7 mm was lower compared to women with an EMT >7 mm (60/258 versus 4981/10354, OR 0.42, 95%CI 0.27 to 0.67) (Kasius et al 2014). The difference between the association of EMT and pregnancy rates in IUI versus IVF might be explained by the difference in hormone treatment. In IVF, a GnRH agonist or GnRH antagonist is used for downregulation of the menstrual cycle, which might influence the development of endometrial lining. In IUI it is not common to use downregulation.

In conclusion, CC leads to a significantly thinner endometrium compared to gonadotrophins, but since there was no evidence of a consistent association between EMT and the ongoing pregnancy rate, a relatively thin endometrium after CC is not a valid reason to prefer gonadotrophins as the stimulation agent in IUI for unexplained subfertility.

ACKNOWLEDGEMENTS

We thank all couples that participated in the SUPER study, the hospitals and their staff, the research nurses and the staff of the Dutch Consortium for Healthcare Evaluation and Research in Obstetrics and Gynaecology for logistic support and the staff of the Clinical Research Unit of the Academic Medical Centre, Amsterdam for their help with the randomization program and the online database.

COMPETING INTERESTS

Prof dr BWJ Mol is supported by a NHMRC Practitioner Fellowship (GNT1082548) and reports consultancy for Merck, ObsEva and Guerbet.

FUNDING

The initial RCT received funding from the Dutch Organisation for Health Research and Development (ZonMw) (Health Care Efficiency Research; projectnumber *80-83600-98-10192)* .

REFERENCES

Calhaz-Jorge C, De Geyter C, Kupka M, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C and Goossens V. European IVF-monitoring Consortium (EIM); European Society of Human Reproduction and Embryology (ESHRE), Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod. 2017; 10:1957-1973

Danhof NA, van Wely M, Repping S, Koks C, Verhoeve HR, de Bruin JP, Verberg MFG et al. Follicle stimulating hormone (FSH) versus Clomiphene Citrate (CC) in intrauterine insemination for unexplained subfertility: A randomized controlled trial. Hum Reprod. 2018; dey268

Gadalla MA, Huang S, Wang R, Norman RJ, Abdullah SA, El Saman AM, Ismail AM, van Wely M, Mol BWJ. Effect of clomiphene citrate on endometrial thickness, ovulation, pregnancy and live birth in anovulatory women: systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2018;51(1):64-76.

Karlsson B, Granberg S, Ridell B, Wikland M. Endometrial thickness as measured by transvaginal sonography: interobserver variation. Ultrasound Obstet Gynecol. 1994;4(4):320-5.

Kasius A, Smit JG, Torrance HL, Eijkemans MJ, Mol BW, Opmeer BC, Broekmans FJ. Endometrial thickness and pregnancy rates after IVF: a systematic review and meta-analysis. Hum Reprod Update. 2014;20(4):530-41.

Kettel LM, Roseff SJ, Berga SL et al. Hypothalamic-pituitary-ovarian-response to clomiphene citrate in women with polycystic ovary syndrome. Fertil Steril. 1993;59:532-8

Massai MR, de Ziegler D, Lesobre V, Bergeron C, Frydman R, Bouchard P. Clomiphene citrate affects cervical mucus and endometrial morphology independently of the changes in plasma hormonal levels induced by multiple follicular recruitment. Fertil Steril. 1993;59(6):1179-86.

Schipper I, Hop WC, Fauser BC. The follicle-stimulating hormone (FSH) threshold/window concept examined by different interventions with exogenous FSH during the follicular phase of the normal menstrual cycle: duration, rather than magnitude, of FSH increase affects follicle development. J Clin Endocrinol Metab. 1998;83(4):1292-8.

The Practice Committee of the American Society for Reproductive Medicine Effectiveness and treatment for unexplained infertility Fertil Steril 2006; 86(4):S111-S114

Unfer V, Costabile L, Gerli S, Papaleo E, Marelli G, Di Renzo GC. Low dose of ethinyl estradiol can reverse the antiestrogenic effects of clomiphene citrate on endometrium. Gynecol Obstet Invest. 2001;51(2):120-3.

Weiss NS, van Vliet MN, Limpens J, Hompes PGA, Lambalk CB, Mochtar MH, van der Veen F, Mol BWJ, van Wely M. Endometrial thickness in women undergoing IUI with ovarian stimulation. How thick is too thin? A systematic review and meta-analysis. Hum Reprod. 2017;32(5):1009-1018.

Wu CH. Plasma hormones in clomiphene citrate therapy. Obstet gyn. 1977;49:443-8

Yeko TR, Nicosia SM, Maroulis GB, Bardawil WA, Dawood MY. Histology of midluteal corpus luteum and endometrium from clomiphene citrate-induced cycles. Fertil Steril. 1992;57(1):28-32.

Zaidi J, Campbell S, Pittrof R, Tan SL. Endometrial thickness, morphology, vascular penetration and velocimetry in predicting implantation in an in vitro fertilization program. Ultrasound Obstet Gynecol. 1995;6(3):191-8.

7.

Gonadotrophins or clomiphene citrate in couples with unexplained infertility undergoing intrauterine insemination: a cost-effectiveness analysis

authors:

N.A. Danhof, Madelon van Wely, S. Repping, D.P. van der Ham, N. Klijn, C.A.H. Janssen, J.M. Rijn -van Weert, M. Twisk, M.A.F. Traas, M.J. Pelinck, D.A.M. Perquin, D.E.S. Boks, A. Sluijmer, B.W.J. Mol, F. van der Veen, M. H. Mochtar

Reproductive BioMedicine Online 2019;19:pii:S1472-6483.30740-0

ABSTRACT

Research question: What is the cost-effectiveness of gonadotrophins compared with clomiphene citrate in couples with unexplained subfertility undergoing intrauterine insemination (IUI) with ovarian stimulation under strict cancellation criteria?

Design: A cost-effectiveness analysis alongside a randomized controlled trial. Between July 2013 and March 2016, we randomized 738 couples to gonadotrophins (369) or clomiphene citrate (369) in a multicentre randomized controlled trial in the Netherlands. We compared the direct medical costs of both strategies. Direct medical costs included costs of medication, cycle monitoring, insemination and, if applicable, pregnancy monitoring. Non-parametric bootstrap resampling was used to investigate the effect of uncertainty in estimates. The cost-effectiveness analysis was performed according to intention-to-treat. The incremental cost-effectiveness ratio (ICER) between gonadotrophins and clomiphene citrate for ongoing pregnancy and live birth was assessed.

Results: The mean costs per couple were €1 534 for gonadotrophins and €1 067 for clomiphene citrate (mean difference of €468 (95% CI, €464 to €472)). As ongoing pregnancy rates were 31% in women allocated to gonadotrophins and 26% in women allocated to clomiphene citrate (relative risk 1.16, 95% CI 0.93 to 1.47), the incremental cost-effectiveness ratio was €21 804 (95% CI, €11 628 to €31 980) per additional ongoing pregnancy with gonadotrophins and €17 044 (95% CI, €8 998 to €25 090) per additional live birth with gonadotrophins.

Conclusions: Gonadotrophins are more expensive compared with clomiphene citrate in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria, without being significantly more effective.

Trial registration: NTR 4057

BACKGROUND

In most countries, the first line treatment for couples diagnosed with unexplained subfertility is intrauterine insemination (IUI) with ovarian stimulation (Calhaz-Jorge et al., 2017, Practice committee ASRM, 2006).

In a recent multicenter randomized controlled trial comparing gonadotrophins to clomiphene citrate in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria, ongoing pregnancy rates were 31% in women allocated to gonadotrophins and 26% in women allocated to clomiphene citrate (RR 1.16, 95% CI 0.93 to 1.47) , while multiple pregnancy rates were 1% and 2% respectively (RR 0.63, 95% CI 0.21 to 1.89) (Danhof et al., 2018).

These data may be taken as an argument to use gonadotrophins as first line medication in IUI. However, it is essential to gain evidence on the effectiveness of fertility treatments in relation to the costs, since healthcare costs have been rising over the years, mainly due to an increase in pharmaceutical spending (Dieleman et al., 2017). It is therefore needed to first determine the value of gonadotrophins compared with clomiphene citrate in relation to the costs, to be able to make well-founded recommendations on which stimulation agent to use. In this study, a cost-effectiveness analysis was performed from a healthcare perspective in couples with unexplained subfertility undergoing IUI with ovarian stimulation by either gonadotrophins or with clomiphene citrate and adherence to strict cancellation criteria.

MATERIALS AND METHODS

Study design

An economic evaluation was conducted alongside the SUPER study (Danhof et al., 2018) (trial registration: NTR 4057). Here, the specifics of the interventions that contributed to the costs of both treatment arms are discussed. Ethical approval for this study was obtained from the Medical Ethical Committee of the Academic Medical Centre and from the Dutch Central Committee on Research involving Human Subjects on 16th April 2013 (CCMO NL 43131-018-13). Couples were scheduled for a maximum of four cycles of IUI with ovarian stimulation comparing a starting dose of 75 IU gonadotrophins with 100 mg clomiphene citrate. Follicular development was monitored by transvaginal ultrasound and cycles were cancelled if more than three follicles with a diameter of ≥14 mm or five follicles with a diameter of ≥12 mm were seen at transvaginal ultrasound. In case of monofollicular growth, the dose of gonadotrophins was increased by 37·5 IU per day or the dose of clomiphene citrate was increased by 50 mg per day in the next cycle. If a cycle was cancelled due to the development of more than three dominant follicles, the dose of gonadotrophins was decreased by 37·5 IU per day or the dose of clomiphene citrate was decreased by 50 mg per day in the next cycle.

Primary effectiveness outcome was ongoing pregnancy defined as a positive heartbeat at or beyond 12 weeks of gestation. A secondary outcome was multiple pregnancy defined as two or more foetuses with a positive heartbeat at or beyond 12 weeks of gestation.

Economic evaluation

A cost-effectiveness analysis was performed from a healthcare perspective. This implies that the focus was on the direct medical costs during IUI with ovarian stimulation and did not take into account indirect medical costs, which represent costs incurred by the patient in the context of the treatment, like transportation costs, costs that were related to cycle cancellation and productivity loss.

Resource use

Data were collected from the individual case record forms of the RCT. For each woman, the following were documented: the dose of the stimulation agent, the duration of stimulation, whether an ovulation trigger was used and whether an insemination was performed within a time horizon of 6 months or until an ongoing pregnancy occurred. If couples switched their treatment to in vitro fertilization or intracytoplasmic sperm injection (IVF/ICSI), resource use was estimated on the basis of previously published data on resource costs for IVF/ICSI (van Tilborg et al., 2017).

Unit costs

Direct medical costs included the costs of medication, cycle monitoring and insemination. The medication costs were based on the total number of units gonadotrophins or milligrams of clomiphene citrate and the costs of the ovulation trigger. Costs were estimated from the Dutch Formulary on Medication for 2017 (Farmacotherapeutisch

Kompas, https://www.farmacotherapeutischkompas.nl). Costs were calculated for cycle monitoring, insemination and pregnancy monitoring until ongoing pregnancy, based on standardized unit costs as retrieved by an expert panel on cost-effectiveness from the Dutch Consortium for Healthcare Evaluation in Obstetrics and Gynaecology (https://zorgevaluatienederland.nl/associations/1). This Consortium supplied unit costs established of an average of two university and one general hospital. Costs were derived for pregnancy monitoring and delivery from a cost analysis of singleton versus twin pregnancies, in which the costs for a singleton and twin pregnancies up until 6 weeks after delivery were estimated (Lukassen et al., 2004). For miscarriage, the costs of expectant management were used (Lemmers et al., 2016). Costs are expressed in euros (€). The most recently available unit prices were used (for 2017) and corrected for inflation or deflation whenever necessary using the consumer pricing index (Statistics 2017).

Statistical analysis

All outcomes were analysed according to intention-to-treat and costs were calculated by multiplying the quantity of resource use and unit costs. Costs were expressed as means per woman and combined with ongoing pregnancy by calculating the incremental cost-effectiveness ratio (ICER), defined as the ratio of the difference in costs and the difference in effectiveness between gonadotrophins and clomiphene citrate.

The difference in mean costs and ICER was expressed with 95% CI, estimated by 1000 bootstrap replications. The ICER was visualized by plotting a cost-effectiveness plane and cost-effectiveness acceptability curves. The reference strategy was ovarian stimulation with clomiphene citrate. Statistical analyses were performed using SPSS Statistics for Windows, Version 24.0 (IBM Corp., Armonk, NY, USA) and Microsoft Excel 2010 for the bootstrapping.

Per protocol and sensitivity analyses

A per protocol analysis was performed, in which the costs per protocol of IUI with gonadotrophins or clomiphene citrate were combined with ongoing pregnancy by calculating the ICER. Two sensitivity analyses were performed; one with live birth as the effectiveness outcome, and one with the higher dose of gonadotrophins of 150 IU, because this is common practice in some countries (Diamond et al., 2015).

RESULTS

Study group

Between July 2013 and March 2016, 369 women were randomly allocated to ovarian stimulation with gonadotrophins and 369 women to ovarian stimulation with clomiphene citrate. As previously reported in Danhof et al., the baseline characteristics, which were female age, primary subfertility, diagnosis of subfertility median duration of subfertility, current smoking status, body mass index and total motile sperm count,

were well balanced between the randomized couples. There were 113 (31%) ongoing pregnancies following ovarian stimulation with gonadotrophins and 97 (26%) ongoing pregnancies following ovarian stimulation with clomiphene citrate (RR 1.16, 95% CI 0.93 to 1.47). Five women (1.4%) had a multiple pregnancy following ovarian stimulation with gonadotrophins and eight women (2.2%) had a multiple pregnancy following ovarian stimulation with clomiphene citrate (RR 0.63, 95% CI 0.21 to 1.89). 105 women (28%) had a live birth following ovarian stimulation with gonadotrophins and 92 women (25%) had a live birth following ovarian stimulation with clomiphene citrate (RR 1.14, 95% CI 0.90-1.45). 32 women (9%) had a miscarriage following ovarian stimulation with gonadotrophins and 31 women (8%) had a miscarriage following ovarian stimulation with clomiphene citrate (RR 1.03, 95% CI 0.64-1.66).

Economic evaluation

Resource use and unit costs

Table 1 lists the mean resource use. The mean number of IUI cycles per couple was similar for both treatment arms. Unit costs are listed in Table 2.

Table 1. Resource use per woman*

	Gonadotrophins (*n*=369)	**Clomiphene citrate (*n*=369)**
Medication		
Gonadotrophins in IU	1783 (1196)	95.12 (349.13)
Clomiphene citrate in mg	59.55 (334.34)	1339 (945)
hCG in IU	12182 (6988)	13306 (6618)
Intervention		
Number of IUI cycles	2.83 (1.48)	2.90 (1.34)
Number of IVF cycles	0.01 (0.07)	0 (0)
Number of ICSI cycles	0 (0.05)	0.01 (0.09)

* Data are mean (SD)

HCG = human chorionic gonadotrophin; ICSI = intracytoplasmic sperm injection; IUI = intrauterine insemination.

Table 2. Unit costs

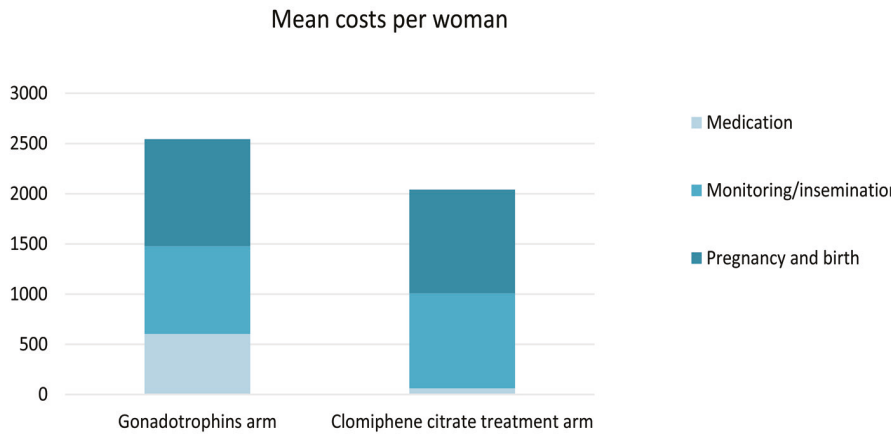
Cost item	Unit	Unit costs (Euros)	Reference
Medication			
Gonadotrophins	75 IU	24.75	Dutch Formulary on medication 2017
Clomiphene citrate	50 mg	0.53	Dutch Formulary on medication 2017
hCG for ovulation induction	5000 IU	5.83	Dutch Formulary on medication 2017
Pregnancy and delivery			
Ongoing pregnancy	1	200.00	Dutch consortium*
Singleton live birth	1	3107.00	Lukassen *et al.* 2004
Multiple live birth	1	16419.00	Lukassen *et al.* 2004
Miscarriage (expectant management)	1	91.00	Dutch consortium*
Intervention			
IUI - monitoring	1	150.54	Dutch consortium*
IUI - insemination	1	170.00	Dutch consortium*
IVF	1	1 365.84	Dutch consortium*
ICSI	1	1 699.13	Dutch consortium*

HCG = human chorionic gonadotropin; ICSI = intracytoplasmic sperm injection; IUI = intrauterine insemination.
* Costs are derived from the expert panel Dutch Consortium for Research in Women's Health

Costs

The mean costs per couple for ongoing pregnancy 6 months after randomization were €1534 for gonadotrophins and €1067 for clomiphene citrate (mean cost difference €468, 95% CI €464–472). The mean costs per couple for live birth 6 months after randomization were €2537 for gonadotrophins and €2042 for clomiphene citrate (mean cost difference €487, 95% CI €473–501) (figure 1).

Figure 1. Mean costs per woman with follow up until live birth

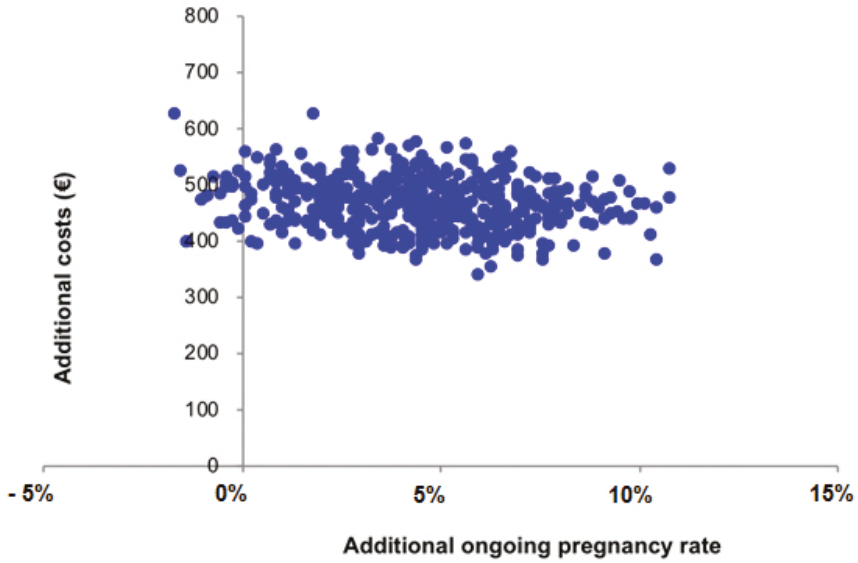


CC=clomiphene citrate

Cost-effectiveness

The ICER for gonadotrophins compared to clomiphene citrate was €21 804 (95%CI, €11 628 to €31 980), reflecting the additional costs to achieve one additional ongoing pregnancy in women allocated to ovarian stimulation with gonadotrophins compared to women allocated to ovarian stimulation with clomiphene citrate. The bootstrap samples were located in both the north western and north eastern quadrant of the plot, with the vast majority in the north eastern quadrant (figure 2). This reflects higher costs for gonadotrophins compared to clomiphene citrate, with slightly more ongoing pregnancies for gonadotrophins compared to clomiphene citrate.

Figure 2. Cost-effectiveness plane of gonadotrophins compared to clomiphene citrate until ongoing pregnancy



Cost-effectiveness plane of gonadotrophins compared with clomiphene citrate until ongoing pregnancy. Each point in the cost-effectiveness plane represents the uncertainty of the additional costs and effect of gonadotrophins compared with clomiphene citrate after non-parametric bootstrap resampling

Per protocol and sensitivity analyses

Of the 738 women, 657 were treated according to protocol and included in the per protocol analysis. 82 (25%) women had an ongoing pregnancy after gonadotrophins compared to 70 (21%) after clomiphene citrate (RR 1.16, 95% CI 1.88 to 1.55). Mean costs per woman were €1 675 for gonadotrophins and €1 078 for clomiphene citrate (mean cost difference €598, 95% CI 595 to 602). The ICER for ovarian stimulation with gonadotrophins compared to clomiphene citrate was €22 782 (95% CI, €10 947 to €34 617) (table 3).

Table 3. Sensitivity analyses

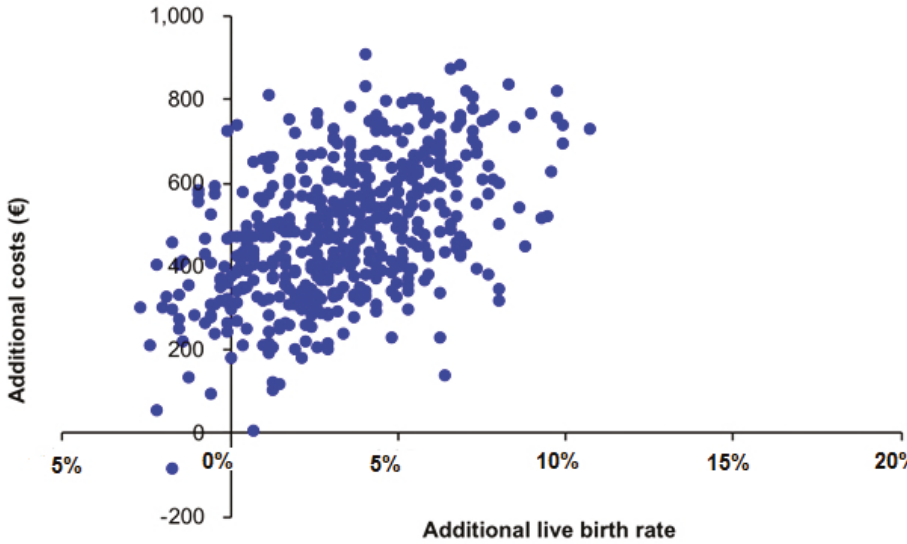
Description	Mean cost gonadotrophins treatment arm	Mean cost clomiphene citrate treatment arm	Mean cost difference (95% CI)	ICER
Per protocol*	1 675.46	1 077.94	598 (595 - 602)	22 782
Live birth as effectiveness outcome**	2 536.88	2 041.82	487 (473 - 501)	17 044
Dosage of 150 IU gonadotrophins***	2 120.90	1 098.61	1 019 (1 013 - 1 024)	42 432

* Based on a per protocol analysis with ongoing pregnancy as effectiveness outcome
** Cost of live birth was included in treatment cost.
*** The dosage of gonadotrophins in gonadotrophins treatment arm was increased, therefore costs were increased. Ongoing pregnancy was effectiveness outcome and remained fixed
ICER=incremental cost-effectiveness rate
All costs are in euros.

With live birth as the effectiveness outcome, the ICER for gonadotrophins compared to clomiphene citrate was €17 044 (95% CI, €8 998 to €25 090) (table 3, figure 3).

When using a dosage of 150 IU gonadotrophins , the ICER for gonadotrophins compared to clomiphene citrate for ongoing pregnancy was €42 432 (95% CI, €25 093 to €59 772) (table 3).

Figure 3. Cost-effectiveness plane of gonadotrophins compared to clomiphene citrate until live birth



Cost-effectiveness plane: gonadotrophins versus clomiphene citrate. Each point in the cost-effectiveness plane represents the uncertainty of the additional costs and effect of FSH compared with clomiphene citrate after nonparametric bootstrap resampling.

DISCUSSION

This cost-effectiveness analysis from a healthcare perspective comparing an ovarian stimulation regimen with gonadotrophins versus clomiphene citrate in couples with unexplained subfertility undergoing four cycles of IUI with adherence to strict cancellation criteria within six months showed that the costs to achieve one additional ongoing pregnancy with gonadotrophins compared to clomiphene citrate were €21 804 and the costs to achieve one additional live birth with gonadotrophins compared to clomiphene citrate were €17 044. Multiple pregnancy rates, which were included in this cost-effectiveness analysis, were low and similar for both stimulation regimens, indicating the good safety profile of a protocol with adherence to strict cancellation criteria, ie cycle cancellation when more than three follicles developed.

This study has several strengths. First, the cost-effectiveness analysis was based on the results of a well-powered multicentre RCT. Second, we presented the incremental costs for the most important effectiveness outcome measures, ie ongoing pregnancy and live birth. In addition to the intention to treat analysis, in which we presented cost effectiveness of clinical daily practice, we also presented the per protocol analysis to provide insight an estimate of the true cost-effectiveness among those who completed the allocated treatment arm. A potential limitation of this study is that we performed the analysis from a healthcare perspective and thus focussed on direct medical costs. We feel it unlikely that including indirect costs from a social perspective, like transportation costs, costs that were related to cycle cancellation and productivity loss, would change our conclusion, as the couples in both treatment arms underwent similar IUI protocols including similar criteria for cycle cancellation; the only difference being the pharmacological agent. Another limitation is that the definition of unexplained subfertility is not internationally established. This has led to a different approach among fertility centres and countries, which might affect the generalizability of our results.

The protocol with 75 IU gonadotrophins as starting dose and adherence to strict cancellation criteria resulted in a multiple pregnancy rate of just 1.4% without compromising the ongoing pregnancy rate. The use of higher dosages of gonadotrophins in IUI is in our view not justified, since this will either lead to a higher cancellation rate or to a higher multiple pregnancy rate for higher costs (Diamond et al., 2015).

The mean costs per couple for IUI with ovarian stimulation were in line with the costs of IUI with ovarian stimulation reported in a cost-effectiveness analysis comparing twelve months of IUI with ovarian stimulation to twelve months of *in vitro* fertilization with single embryo transfer (IVF-SET) (Tjon-Kon-Fat et al., 2015).

In a system in which infertility treatments are reimbursed from public resources, the question arises as to whether €21,804 should be spent to achieve one extra ongoing pregnancy or €17,044 should be spent to achieve one extra live birth. There is no consensus on acceptable costs for one extra ongoing pregnancy or one extra live birth. The NICE fertility guideline recommends a threshold of £30 000 or €32 700 per quality adjusted life year (QALY), but also emphasizes that a QALY is not a proper measuring instrument for an extra ongoing pregnancy or live birth, since QALYs are anticipated to assess health improvements in patients.(NICE 2013) According to a patient preference study women are willing to pay €100 to €500 extra for a few percent increase in pregnancy rates (Palumbo et al., 2011). In a system in which fertility care is private practice, it seems prudent to communicate our data to the patients

In conclusion, this study provides insight in the actual costs of two frequently used stimulation regimens worldwide. Because healthcare costs have been rising over the years, it is essential to be aware of the costs in clinical decision-making. Gonadotrophins are more expensive than clomiphene citrate in couples with unexplained subfertility undergoing IUI with adherence to strict cancellation criteria, without being significantly more effective.

ACKNOWLEDGEMENTS

We thank all couples that participated in the SUPER study, the hospitals and their staff, the research nurses and the staff of the Dutch Consortium for Healthcare Evaluation and Research in Obstetrics and Gynaecology for logistic support and the staff of the Clinical Research Unit of the Academic Medical Centre, Amsterdam for their help with the randomization program and the online database.

COMPETING INTERESTS

Prof dr BWJ Mol is supported by a NHMRC Practitioner Fellowship (GNT1082548). BWM reports consultancy for Merck, ObsEva and Guerbet.

The other authors declare that they have no competing interests.

FUNDING

The initial RCT received funding from the Dutch Organisation for Health Research and Development (ZonMw) (Health Care Efficiency Research; projectnumber *80-83600-98-10192)* .

REFERENCES

Calhaz-Jorge C, De Geyter C, Kupka M, de Mouzon J, Erb K, Mocanu E, Motrenko T, Scaravelli G, Wyns C and Goossens V. European IVF-monitoring Consortium (EIM); European Society of Human Reproduction and Embryology (ESHRE), Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod. 2017, 10:1957-1973

Danhof NA, van Wely M, Repping S, Koks C, Verhoeve HR, de Bruin JP, Verberg MFG et al. Follicle stimulating hormone (FSH) versus Clomiphene Citrate (CC) in intrauterine insemination for unexplained subfertility: A randomized controlled trial. Human Reproduction 2018 oct; 33(10):1866–1874

Diamond MP, Richar SL, Coutifaris C, Alvero R, Robinson R, Casson P et al. Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility. N Engl J Med 2015; 373

Dieleman JL, Squires E, Bui AL, Campbell M, Chapin A, Hamavid H et al. Factors Associated With Increases in US Health Care Spending. JAMA. 2017;318(17):1668-1678

Farmacotherapeutisch kompas. https://www.farmacotherapeutischkompas.nl accessed October 17, 2017)

Lemmers M, Verschoor MA, Oude Rengerink K, Naaktgeboren C, Opmeer BC, Bossuyt PM, Huirne JA, Janssen CA, Radder C, Klinkert ER, et al. MisoREST study group. Surgical versus expectant management in women with an incomplete evacuation of the uterus after misoprostol treatment for miscarriage: a randomized controlled trial. MisoREST: Hum Reprod. 2016, 11:2421-242

Lukassen HG, Schönbeck Y, Adang EM, Braat DD, Zielhuis GA, Kremer JA. Cost analysis of singleton versus twin pregnancies after in vitro fertilization. Fertil Steril. 2004 May;81(5):1240-6

NICE. Fertility: Assessment and Treatment for People with Fertility Problems. 2013.

Palumbo, A., P. De La Fuente, M. Rodriguez, F. Sanchez, J. Martinez-Salazar, M. Munoz, J. Marqueta, J. Hernandez, O. Espallardo, C. Polanco, S. Paz, and L. Lizan. 'Willingness to pay and conjoint analysis to determine women's preferences for ovarian stimulating hormones in the treatment of infertility in Spain', Hum Reprod, 2011 26: 1790-8.

Practice Committee of the American Society for Reproductive Medicine Effectiveness and treatment for unexplained infertility Fertil Steril, 86 (suppl 4) (2006), pp. S111-S114

Tjon-Kon-Fat RI, Bensdorp AJ, Bossuyt PM, Koks C, Oosterhuis GJ, Hoek A, Hompes P, Broekmans FJ, Verhoeve HR, de Bruin JP, et al. Is IVF-served two different ways-more cost-effective than IUI with controlled ovarian hyperstimulation? Hum Reprod. 2015, 100:2331-9

van Tilborg TC, Oudshoorn SC, Eijkemans, MJC, Mochtar MH, van Golde RJT, Hoek A, Kuchenbecker WKH, Fleischer K, de Bruin JP, Groen H, et al. Optimist study group. Individualized FSH dosing based on ovarian reserve testing in women starting IVF/ICSI: a multicentre trial and cost-effectiveness analysis. Hum Reprod 2017, 32: 2485-95

8.

General discussion

In this chapter we weigh the value of intrauterine insemination (IUI) with ovarian stimulation and in vitro fertilisation (IVF) to each other and to expectant management in couples with unexplained subfertility. We then focus on our findings towards an effective and safe ovarian stimulation protocol for IUI in couples with unexplained subfertility and we end our discussion with providing implications for clinical practice and suggestions for further research.

The value of intrauterine insemination with ovarian stimulation and in vitro fertilisation in unexplained subfertility

Forty to fifty percent of all couples seeking medical help because of subfertility will be diagnosed with unexplained subfertility, defined as a failure to conceive after one year of regular unprotected intercourse, a normal semen analysis, an ovulatory cycle and at least one sided tubal patency (Evers, 2002). Intrauterine insemination (IUI) with ovarian stimulation and in vitro fertilization (IVF) are common interventions for couples with unexplained subfertility (ASRM 2006; NVOG 2010; NICE 2017; RCOG 2019).

IUI with ovarian stimulation is a minimal invasive and low-cost treatment, but the multiple pregnancy rate is reported as high (Calhaz-Jorge *et al.*, 2017). Multiple pregnancies are associated with an increased risk of maternal morbidity, neonatal morbidity and neonatal mortality (Ombelet et al 2006). As such, this constitutes a major problem in daily clinical practice. IUI with ovarian stimulation is the major provider of triplets (Braat et al., 2003).

IVF does not carry an increased risk of multiple pregnancies when it is performed with single embryo transfer (SET) instead of double embryo transfer. Effectiveness and safety of single embryo transfer have now convincingly been shown (Mclernon et al., 2010, Maheshwari et al., 2011, Cutting 2018). IVF is obviously more invasive and more costly than IUI with ovarian stimulation.

The recommendation of most clinical guidelines for the management of couples with unexplained subfertility is to start an intervention after one or two years of unprotected intercourse without conceiving and to begin with the least invasive intervention before moving on to a more invasive one (ASRM 2006; NICE 2017; NVOG 2010; RCOG 2019). The main issue here is that this recommendation is not based on sound evidence taking effectiveness and safety into account. This is due to scarcity of randomized controlled trials that compare IUI with ovarian stimulation or IVF to expectant management in head to head comparisons in couples with unexplained subfertility.

To address this, in an attempt to move the field further, we performed a network meta-analysis including randomized controlled trials on IUI with ovarian stimulation, IVF and expectant management in couples with unexplained subfertility and assessed their effectiveness and safety (Chapter 2). Network meta-analysis compares multiple treatments simultaneously by using both direct and indirect evidence and provides a hierarchy of these treatments, which can potentially better inform clinical decision-making.

Concerning effectiveness, the network meta-analysis showed that IVF -in any kind of protocol reported- followed by IUI with ovarian stimulation -in any kind of protocol reported- is more likely to result in more live births than expectant management. These outcomes represent the best available evidence so far, which is -and we would like to emphasize this- of low to moderate quality. This is mainly due to the fact that apart from the comparison IUI with ovarian stimulation and IVF, other comparisons have relatively few included studies with direct evidence, which means that a large part of the final estimates is based on indirect evidence. Since the effectiveness of fertility treatments is dependent on the prognosis of natural conception, we also analysed the effect of the duration of subfertility –which determines prognosis to a large extent. This subgroup analysis showed that IUI with ovarian stimulation and IVF increased the chance of live births or ongoing pregnancies compared to expectant management, but limited to couples with a duration of subfertility more than two years. These results should be interpreted with caution, since this subgroup analysis was based on different mean/median values. Nevertheless, this finding is in line with a previous cohort study, in which IUI with ovarian stimulation was associated with higher chances of ongoing pregnancy compared to expectant management, specifically in couples with poor prognoses of natural conception, i.e. <15% over 6 months or <25% over 1 year (van Eekelen et al., 2019).

Concerning safety, the network meta-analysis showed that IUI with ovarian stimulation -in any kind of protocol reported- and IVF -in any kind of protocol reported- resulted in more multiple pregnancies than expectant management. The explanation for this is probably the variability between the protocols in which IUI with ovarian stimulation and IVF were executed. This pinpoints to the relevance of improving safety in medically assisted reproduction.

The second part of the thesis was dedicated to find a balance between effectiveness and safety of IUI with ovarian stimulation.

A safe and effective ovarian stimulation protocol for IUI

Multiple pregnancies after fertility treatments are nowadays unacceptable and therefore it is essential that IUI with ovarian stimulation is not only effective, but also safe. To reduce the risk of multiple pregnancies, IUI with ovarian stimulation should not aim for more than two follicles per cycle (Van Rumste et al., 2008). In a randomized controlled trial comparing ovarian stimulation with gonadotrophins to clomiphene citrate in IUI for unexplained subfertility, this was put into a clinical context by limiting the number of dominant follicles per cycle through withholding insemination when more than three dominant follicles developed (Dankert et al., 2007). The result was a low multiple pregnancy rate without compromising effectiveness for both ovarian stimulation with gonadotrophins and clomiphene citrate. Nevertheless, uncertainty remained as this study was underpowered. We therefore evaluated the effectiveness and safety of a protocol with adherence to strict cancellation criteria in IUI with ovarian stimulation and established which pharmaceutical agent should be the drug of choice.

We also investigated whether individualized treatment selection markers at baseline could aid in identifying couples who would have better chances of an ongoing pregnancy with IUI with ovarian stimulation with one or the other pharmaceutical agent and we evaluated costs.

In chapter 3 we evaluated the safety of gonadotrophins and clomiphene citrate in IUI within a protocol with adherence to strict cancellation criteria in couples with unexplained subfertility and a poor prognosis of natural conception in a randomized controlled trial (SUPER study). The SUPER study showed similar and acceptable low multiple pregnancy rates for both gonadotrophins and clomiphene citrate. On the other hand, gonadotrophins resulted in a slightly higher, but statistically insignificant, ongoing pregnancy rate and live birth rate compared to clomiphene citrate. To increase statistical power, we next conducted a network meta-analysis comparing pharmaceutical agents for ovarian stimulation in IUI with adherence to strict cancellation criteria, in which we included the SUPER study and the only two other randomized controlled trials comparing gonadotrophins and clomiphene citrate in IUI within a protocol with adherence to strict cancellation criteria. The end result was that IUI with gonadotrophins and adherence to strict cancellation criteria resulted in a higher, albeit non-significant, ongoing pregnancy rate and live birth rate compared to IUI with clomiphene citrate and adherence to strict cancellation criteria in couples with unexplained subfertility (gonadotrophins 22% vs clomiphene citrate 18%, RR 1.20, 95% CI 0.95 to 1.51). Both gonadotrophins and clomiphene citrate resulted in acceptable multiple pregnancy rates, based on a pairwise meta-analysis (gonadotrophins 2% vs clomiphene citrate 2%, RR 0.80, 95% CI 0.38 to 1.68). We were not able to perform network meta-analysis for the primary safety outcome multiple pregnancy rate because of high heterogeneity. We have endeavoured to present the evidence as complete as possible, since we reported a pairwise meta-analysis and a subgroup analysis including studies that performed IUI with adherence to strict cancellation criteria. Therefore, we feel that the evidence regarding to gonadotrophins and clomiphene citrate in relation to multiple pregnancies is robust (Chapter 4).

The data presented in chapters 3 and 4 clearly show that the multiple pregnancy rate is low within a protocol with adherence to strict cancellation criteria without compromising effectiveness, but still the majority of the couples did not conceive after IUI with ovarian stimulation. This raised the question whether the ongoing pregnancy rates would have been higher following gonadotrophins or CC for some specific couples, ie whether a more individualized approach would be possible. We therefore aimed to identify whether individualized treatment selection markers at baseline could aid in identifying couples who would have better chances of an ongoing pregnancy with IUI with ovarian stimulation with one or the other pharmaceutical agent. A secondary analysis of the SUPER study assessed female age, body mass index, duration of subfertility, primary or secondary subfertility and total motile count as potential treatment selection markers. This analysis was unable to identify any

treatment selection markers at baseline. Since our sample size was rather small for a secondary analysis, it was impossible to detect smaller associations between potential treatment selection markers and ongoing pregnancy, although the clinical relevance of any smaller associations might be questioned (Chapter 5). Since pharmaceutical agents for ovarian stimulation do not only affect folliculogenesis but also have an effect on endometrial thickness, we analyzed the value of the impact of gonadotrophins and clomiphene citrate on endometrial thickness and on ongoing pregnancy rates in another secondary analysis of the SUPER study. If this were the case, this would provide an extra tool to improve effectiveness. Gonadotrophins lead to a 1.4mm thicker endometrium compared to clomiphene citrate. This difference in endometrial thickness between gonadotrophins and clomiphene citrate did not affect ongoing pregnancy rates and therefore endometrial thickness is not a reason to cancel cycles. Although we cannot exclude that a difference of 1.4 mm in endometrial thickness is due to interobserver variability or due to differences in equipment, the impact of this is most likely to be limited, since our study was based on a multicentre randomized controlled trial with women from 24 clinics, thereby reflecting daily clinical practice (Chapter 6)

The cost-effectiveness analysis from a healthcare perspective in couples with unexplained subfertility based on the SUPER study identified an incremental cost-effectiveness ratio of €21 804 (95% CI, €11 628 to €31 980) per additional ongoing pregnancy with gonadotrophins and €17 044 (95% CI, €8 998 to €25 090) per additional live birth with gonadotrophins compared to clomiphene citrate (Chapter 7). In a system in which fertility care is private practice, it seems careful to communicate these data to the patients. In a system in which infertility treatments are reimbursed from public resources, the question rises whether €21 804 should be spent to achieve one extra ongoing pregnancy or €17 044 should be spent to achieve one extra live birth. From a public resource perspective and taking an arbitrary amount of €100 000 to spend, an estimated 18 to 23 ongoing pregnancies can be achieved with gonadotrophins and 22 to 30 ongoing pregnancies can be achieved with CC. This warrants a careful weighing of how to spend public resources in a system where money is limited.

Based on the data obtained by the studies described in this part of the thesis, we conclude that in couples with unexplained subfertility and a poor prognosis of natural conception, IUI regimens with adherence to strict cancellation criteria lead to acceptable multiple pregnancy rates without compromising effectiveness. We could not identify specific subgroups of women at baseline that benefit from one or the other pharmaceutical agent and also endometrial thickness is not a valid reason to switch between pharmaceutical agents. Within a strategy with adherence to strict cancellation criteria, we therefore would suggest to perform IUI with gonadotrophins, if costs do not play a role. Otherwise, we would suggest to start with IUI with clomiphene citrate.

Implications for clinical practice

We recommend IUI with ovarian stimulation and adherence to strict cancellation criteria as first line treatment in couples with unexplained subfertility and a prognosis of natural conception below 30%, since it is more effective than expectant management and more cost-effective, equally safe and less invasive compared to IVF-SET.

Hopefully, this regimen will soon be implemented in clinical practice to avoid high risks for multiple pregnancies, regardless of the setting in which reproductive services are provided. We would suggest to develop an international guideline on how to perform IUI for unexplained subfertility.

Gonadotrophins are more effective, but also more costly, compared to clomiphene citrate. This may be important knowledge in any type of shared decision making, with the caveat that resources in a health care system based on solidarity are not endless.

Implications for further research

Based upon the results presented in this thesis we have the following recommendations for further research.

Our first recommendation is to confirm the cost-effectiveness of IUI with ovarian stimulation and IVF-SET compared to expectant management in couples with unexplained subfertility in randomized controlled trials, because the randomized controlled trials included in our network meta-analysis were of low to moderate quality evidence.

Second, the greatest benefit of IUI with ovarian stimulation is in couples with a poor prognosis of natural conception. To date, the boundary between intermediate prognosis and poor prognosis is set on 30%. Whether this is the correct boundary level is unknown. Hence studies comparing IUI with ovarian stimulation to expectant management –"dose finding studies"- in couples with various prognosis is still needed.

Third, in this thesis we provided evidence on first line treatment for couples with unexplained subfertility and a poor prognosis of natural conception, but evidence on the most effective second line treatment is lacking. More research is needed among couples with unexplained subfertility that did not conceive after six to nine cycles of IUI with ovarian stimulation. Is in this population IVF-SET the best next step to take? This could be investigated in large cohort studies or in randomized controlled trials. Couples who did not succeeded in conceiving following six to nine cycles of IUI could be included. They could then be randomized between continuing IUI or start with IVF.

Fourth, since we were unable to find prognostic factors identifying couples at baseline that would benefit from one or the other pharmaceutical agent and to determine any association between the impact of gonadotrophins and clomiphene citrate on endometrial thickness and ongoing pregnancy, we suggest to gather individual patient data of the studies included in the network meta-analysis to confirm or refute our

findings, taking prognosis on natural conception into account. The same could then be done for the cost-effectiveness analysis.

Fifth, in our network meta-analysis on pharmaceutical agents for ovarian stimulation in IUI we also included Letrozole, but randomized controlled trials that performed IUI with Letrozole within a protocol with adherence to strict cancellation criteria are lacking. Since the latter may be key, we suggest a randomized controlled trial comparing Letrozole to clomiphene citrate in IUI with strict cancellation criteria in a superiority design with an economic evaluation from a healthcare perspective alongside it to determine the most cost-effective and safe ovarian stimulation regimen in IUI for unexplained subfertility.

Finally, the cause of unexplained infertility is still an enigma and clinicians are poorly informed about the sexual health of these couples. More research is needed in the sexual health of subfertile couples and whether this influences the pregnancy chances.

REFERENCES

Braat DD, Schönbeck Y, Kremer JA. Multiple pregnancies: epidemiology and management. Ned Tijdschr Geneeskd. 2003 Oct 4;147(40):1952-5

Calhaz-Jorge C, De Geyter C, Kupka MS, de Mouzon J, Erb K, Mocanu E, et al. Assisted reproductive technology in Europe, 2013: results generated from European registers by ESHRE. Hum Reprod. 2017;32:1957-73.

Cutting R. Single embryo transfer for all. Best Pract Res Clin Obstet Gynaecol. 2018 Nov;53:30-37.

Dankert T, Kremer JA, Cohlen BJ, Hamilton CJ, Pasker-de Jong PC, Straatman H and van Dop PA. A randomized clinical trial of clomifene citrate versus low dose recombinant FSH for ovarian hyperstimulation in intrauterine insemination cycles for unexplained and male subfertility. Hum Reprod. 2007;3:792-7

Evers JL. Female subfertility. *Lancet* 2002;360:151-159.

Maheshwari A, Griffiths S, Bhattacharya S. Global variations in the uptake of single embryo transfer. Hum Reprod Update. 2011 Jan-Feb;17(1):107-20.

McLernon DJ, Harrild K, Bergh C, Davies MJ, de Neubourg D, Dumoulin JC, Gerris J, Kremer JA, Martikainen H, Mol BW, Norman RJ, Thurin-Kjellberg A, Tiitinen A, van Montfoort AP, van Peperstraten AM, Van Royen E, Bhattacharya S. Clinical effectiveness of elective single versus double embryo transfer: meta-analysis of individual patient data from randomised trials. BMJ. 2010 Dec 21;341:c6945

NICE guideline 2017, https://www.nice.org.uk/guidance/cg156/chapter/Recommendation

NVOG guideline: onverklaarde subfertiliteit, 2010, https://www.nvog.nl/wp-content/uploads/2017/12/Onverklaarde-subfertiliteit-1.0-15-09-2010.pdf

Ombelet W, Martens G, De Sutter P, Gerris J, Bosmans E, Ruyssinck G, et al. Perinatal outcome of 12,021 singleton and 3108 twin births after non-IVF-assisted reproduction: a cohort study. Hum Reprod. 2006;21(4):1025-32.

Practice Committee of the American Society for Reproductive Medicine. Effectiveness and treatment for unexplained infertility. Fertil steril. 2006;86(5):111-114

Practice committee of the American Society of Reproductive Medicine. Multiple gestation associated with infertility therapy: an American Society for Reproductive Medicine Practice Committee opinion. Fertil Steril. 2012;97(4):825-34

RCOG 2019, https://www.rcog.org.uk/en/patients/fertility/treatment/iui/

van Eekelen R, van Geloven N, van Wely M, McLernon DJ, Mol F, Custers IM, Steures P, Bhattacharya S, Mol BW, van der Veen F, Eijkemans MJ. Is IUI with ovarian stimulation effective in couples with unexplained subfertility? Hum Reprod. 2019 Jan 1;34(1):84-91

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 Nov-Dec;14(6):563-70.

9.

Summary

For countless couples, reproduction is an important aspect of their lives and can be considered as a basic human need. Suffering from subfertility can be of major burden. Forty to fifty percent of all couples seeking medical help because of subfertility will be diagnosed with unexplained subfertility (Evers, 2002). There are many crucial unanswered questions about unexplained subfertility. Since the pathophysiological mechanism of "unexplained subfertility" is unknown, there is no causal treatment for this condition. By lack of anything better, pragmatic treatment options like intrauterine insemination (IUI) with ovarian stimulation and in vitro fertilization (IVF), have been introduced in clinical practice.

The work presented in this thesis first focused on the effectiveness and safety of IUI with ovarian stimulation, in vitro fertilization (IVF) compared to expectant management in couples with unexplained subfertility. Second, we investigated the effectiveness and safety of pharmaceutical agents in IUI within a strategy with adherence to strict cancellation criteria to reduce the risk of multiple pregnancies in couples with unexplained subfertility and a poor prognosis of natural conception.

In **chapter 1** we provided a general introduction and outline of this thesis.

In **chapter 2** we used network meta-analysis to compare the effectiveness and safety of ovarian stimulation (OS) combined with intercourse, intrauterine insemination (IUI) without or with ovarian stimulation and in vitro fertilization (IVF) or intracytoplasmic sperm injection (ICSI) to each other and to expectant management in couples with unexplained subfertility. We compared all these treatment options by a network meta-analysis. We found 28 randomised controlled trials comparing these treatments with each other in a total of 4 469 couples with unexplained subfertility. There was no conclusive evidence of a difference between OS and expectant management (OR 1.09, 95% CI 0.54 to 2.23, mixed evidence, low quality of evidence) or IUI and expectant management (OR 1.45, 95% CI 0.71 to 2.97, mixed evidence, low quality of evidence). Compared to expectant management, the odd ratio of live birth of OS- IUI was 1.82, (95% CI 0.98 to 3.39, mixed evidence, low quality of evidence). The odds ratio of live birth of IVF/ICSI compared to expectant management was 2.87 (95% CI 1.31 to 6.29, mixed evidence, low quality of evidence). These suggest that if the chance of live birth/ ongoing pregnancy following expectant management is assumed to be 15.9%, the chance following OS, IUI, OS-IUI, IVF/ICSI would be 9.3%-29.6%, 11.8%-35.9%, 15.6%-39.0%, and 19.8%-54.3%, respectively. The surface under the cumulative ranking curve values for the expectant management, OS, IUI, OS-IUI and IVF/ICSI were 14.9%, 24.4%, 45.5%, 67.3%, 97.9%, respectively. This suggests that IVF/ICSI is the most effective intervention in terms of live birth or ongoing pregnancy, followed by OS-IUI, IUI, OS and expectant management.

Subgroup analysis showed that in couples with a duration of infertility <2 years, there was no conclusive evidence of any benefit of OS- IUI (OR 0.82, 95% CI 0.45 to 1.49, mixed evidence, low quality of evidence) or IVF/ICSI over expectant management (OR 1.05,

95% CI 0.46 to 2.43, indirect evidence, low quality of evidence). In couples with a longer duration of infertility >2 years, both IVF/ ICSI (OR 4.60, 95% CI 1.74 to 12.18, mixed evidence, moderate quality of evidence) and OS-IUI resulted in more live births/ ongoing pregnancy (OR 2.84, 95% CI 1.25 to 6.43, mixed evidence, low quality of evidence) compared to expectant management.

Compared to expectant management, OS-IUI and IVF/ICSI increased the odds of multiple pregnancy (OR 2.75, 95% CI 0.82 to 9.15, mixed evidence, low quality of evidence; OR 2.59, 95% CI 0.69 to 9.67, mixed evidence, low quality of evidence). These suggest that if the chance of multiple pregnancy following IUI is assumed to be 0.5%, the chance following OS-IUI and IVF/ICSI would be 0.4%-4.2%, and 0.3%-4.5%, respectively. The surface under the cumulative ranking curve values for the expectant management, OS, IUI, OS-IUI and IVF/ICSI were 70.7%, 43.2%, 94.0%, 17.6%, 24.5%, respectively. This suggests that IUI and expectant management are the safest intervention in terms of less multiple pregnancies, followed by OS, IVF/ICSI and OS-IUI.

The evidence was overall of low to moderate quality. Downgrading the level of evidence is mainly due to the fact that apart from the comparison of IUI with ovarian stimulation and IVF, comparisons have relatively few included studies with direct evidence, which means that most results are based on indirect evidence. Therefore, outcomes should be handled with caution. In conclusion, in couples with unexplained subfertility, IVF/ ICSI and OS-IUI may result in more live births/ongoing pregnancies, but also in more multiple pregnancies in comparison with expectant management. IVF/ICSI and OS-IUI are the most effective interventions for live birth/ongoing pregnancy.

In **chapter 3** we presented the results of a randomised controlled trial on the effectiveness of gonadotrophins compared to clomiphene citrate (CC) in couples with unexplained subfertility undergoing intra uterine insemination for a maximum of four cycles within six months. The quintessence of this study was the adherence to strict cancellation criteria to reduce the risk of multiple pregnancies. The primary outcome was the ongoing pregnancy rate and an important secondary outcome was the multiple pregnancy rate. 738 couples were randomized to gonadotrophins (n=369) or to CC (n=369). There were 113 ongoing pregnancies (31%) in the gonadotrophins treatment arm and 97 ongoing pregnancies (26%) in the CC treatment arm (RR 1.16, 95% CI 0.93 to 1.47). There were 5 multiple pregnancies (1%) in the gonadotrophins treatment arm and 8 multiple pregnancies (2%) in the CC treatment arm (RR 0.63, 95% CI 0.21 to 1.89). This led us to the conclusion that adherence to strict cancellation criteria is a successful solution for multiple pregnancy in IUI and that pregnancy outcomes were comparable between gonadotrophins and clomiphene citrate.

In **chapter 4** we used network analysis to compare the effectiveness and safety of IUI with clomiphene citrate, Letrozole or gonadotrophins and natural cycle IUI. We systematically searched the literature on randomized controlled trials that compared any stimulation regimen with CC, Letrozole or gonadotrophins to each other or to

natural cycle IUI among couples with unexplained infertility and were able to include 26 eligible studies reporting on 5316 women. The relative risk on live birth/ongoing pregnancy rates comparing gonadotrophins to CC was 1.39 (95% CI 1.09-1.76, moderate certainty of evidence), comparing Letrozole to CC was 1.09 (95% CI 0.76-1.57, moderate certainty of evidence) and comparing Letrozole to gonadotrophins was 0.79 (95% CI 0.54-1.15, moderate certainty of evidence). We did not perform network meta-analysis on multiple pregnancy due to high inconsistency. Pairwise meta-analyses showed a relative risk on multiple pregnancy rates of 9.11 (95% CI 1.18-70.32) comparing IUI with gonadotrophins to natural cycle IUI. There was no data available on multiple pregnancy rates following IUI with CC or Letrozole compared to natural cycle IUI. The relative risk on multiple pregnancy rates comparing gonadotrophins to CC was 1.42 (95% CI 0.68-2.97), comparing Letrozole to CC was 0.97 (95% CI 0.47-2.01) and comparing Letrozole to gonadotrophins was 0.29 (95% CI 0.14-0.58). In a meta-analysis among studies with adherence to strict cancellation criteria, the relative risk of live births/ ongoing pregnancy rates comparing gonadotrophins to CC was 1.20 (95% CI 0.95-1.51) and the relative risk on multiple pregnancy rates comparing gonadotropins to CC was 0.80 (95% CI 0.38-1.68).

In conclusion, although IUI with gonadotrophins ranked highest on live birth/ongoing pregnancy rates, women undergoing this treatment protocol were also at risk for multiple pregnancies with high complication rates. IUI regimens with adherence to strict cancellation criteria lead to an acceptable multiple pregnancy rate without compromising the effectiveness. Within a protocol with adherence to strict cancellation criteria, gonadotrophins seem to improve live birth/ongoing pregnancy rates compared to CC.

In **chapter 5** we aimed to identify whether individualised treatment selection markers could aid in identifying couples who would have better chances of an ongoing pregnancy with IUI with one or the other pharmaceutical agent. We performed a secondary analysis of the SUPER study, of which we described the essentials in chapter 3. As potential treatment selection markers, we assessed female age, body mass index (BMI), duration of subfertility, whether a couple was primary or secondary subfertile and total motile count (TMC). For each of these factors we developed a logistic regression model to assess if within strata of the factor there were different estimated effects of gonadotrophins versus CC. None of the treatment selection markers were associated with better ongoing pregnancy chances after IUI with gonadotrophins compared to CC (p ranging from 0.123-0.837), so we could not identify treatment selection markers at baseline in couples with unexplained subfertility to choose for ovarian stimulation with gonadotrophins instead CC in IUI. Therefore, based on these data, we could not identify specific groups that have better chances of an ongoing pregnancy with gonadotrophins instead of CC.

In **chapter 6** we evaluated the impact of gonadotrophins and CC on endometrial thickness (EMT) and the association between EMT and ongoing pregnancies. We

therefore performed a secondary analysis on the randomized controlled trial presented in chapter 3. We determined the difference in EMT between women randomised to gonadotrophins and CC over all cycles using a linear mixed model. We investigated the association between EMT and ongoing pregnancy after IUI using a logistic regression model, adjusted for known pregnancy predictors. 666 couples underwent 1968 IUI cycles in total. The mean EMT was 8.9 mm (SD 2.1) in women treated with gonadotrophins and 7.5 mm (SD 2.1) in women treated with CC (Mean difference 1.4 mm; 95% CI: 1.1 to 1.7). The mean EMT was 8.4 mm (SD 2.2) in women that conceived leading to ongoing pregnancy and 8.2 mm (SD 2.2) in women that not conceived (Mean difference 0.29 (-0.10 to 0.68). We found no clear association between EMT and ongoing pregnancy (odds ratio: 1.07 per 1 mm increase, 95% CI 0.99 to 1.15). In women undergoing IUI with ovarian stimulation, a thin EMT after CC is therefore not a valid reason to switch stimulation agents.

In **chapter 7** we described the results of a cost-effectiveness analysis alongside the randomized controlled trial presented in chapter 3. This study was performed from a health care perspective. We compared the direct medical costs of both strategies. Direct medical costs included the costs of medication, cycle monitoring, insemination and, if applicable, pregnancy monitoring. We collected the resource use from case record forms and estimated costs from public sources and published literature. We assessed the mean costs per couple and the incremental cost-effectiveness ratio (ICER) between gonadotrophins and CC. The mean costs per couple were €1 534 for gonadotrophins and €1 067 for CC (mean difference of €468 (95% CI, €464 to €472)). The incremental cost-effectiveness ratio was €21 804 (95% CI, €11 628 to €31 980) per additional ongoing pregnancy with gonadotrophins. In a system in which fertility care is private practice, it seems careful to communicate these data to the patients. In a system in which infertility treatments are reimbursed from public resources, the question rises whether €21 804 should be spent to achieve one extra ongoing pregnancy. This is obviously an issue to be addressed by health care buyers.

REFERENCE

Evers JL. Female subfertility. *Lancet* 2002;360:151-159.

10.

Samenvatting

SAMENVATTING

Voor veel paren vormt voortplanting een belangrijk onderdeel van hun leven. Het kan dan ook beschouwd worden als een menselijke basisbehoefte. Subfertiliteit kan daarom als een grote last worden ervaren. Bij ongeveer veertig tot vijftig procent van alle paren die medische ondersteuning zoeken vanwege subfertiliteit wordt de diagnose "onverklaarde subfertiliteit" gesteld (Evers, 2002). Veel fundamentele vragen over onverklaarde subfertiliteit wachten nog op een antwoord. Omdat het pathofysiologische mechanisme van "onverklaarde subfertiliteit" onbekend is, bestaat er geen causale therapie voor deze aandoening. Bij gebrek aan betere opties, en wellicht ook omdat het nu eenmaal de binnen het werkveld beschikbare middelen zijn, worden pragmatische behandelmethodes zoals intra-uteriene inseminatie (IUI) met ovariële stimulatie en in-vitro fertilisatie (IVF) aangeboden aan paren met onverklaarde subfertiliteit.

In ons onderzoek hebben wij ons eerst gericht op de effectiviteit en veiligheid van IUI met ovariële stimulatie en in-vitro fertilisatie (IVF) in vergelijking met afwachtend beleid voor paren met onverklaarde subfertiliteit. Vervolgens hebben wij onderzocht op welke manier IUI met ovariële stimulatie het meest effectief en veilig kan worden uitgevoerd met een strategie waarbij strikte cancelcriteria worden nageleefd om het risico op een meerlingzwangerschap te verkleinen bij paren met onverklaarde subfertiliteit.

In **hoofdstuk 1** geven wij een algemene introductie en beschrijven wij de doelstellingen van dit proefschrift.

In **hoofdstuk 2** beschrijven wij de resultaten van een netwerk meta-analyse waarin de effectiviteit en veiligheid van ovariële stimulatie met coitus (OS-C), intra uteriene inseminatie (IUI) met of zonder ovariële stimulatie en in vitro fertilisatie (IVF) of intra cytoplasmatische sperma injectie (ICSI) wordt vergeleken met elkaar en met een afwachtend beleid in paren met onverklaarde subfertiliteit. Wij vergeleken deze behandeling in een netwerk meta-analyse. We includeerden 28 gerandomiseerde studies (4 469 paren). Er was geen conclusief bewijs van een verschil tussen OS en een afwachtend beleid (OR 1.09, 95% BI 0.54 tot 2.23, gemixt bewijs, lage kwaliteit) of IUI en een afwachtend beleid (OR 1.45, 95% BI 0.71 tot 2.97, gemixt bewijs, lage kwaliteit). Vergeleken met een afwachtend beleid was de kans op een levend geborene/doorgaande zwangerschap hoger bij OS-IUI (OR 1.82, 95% BI 0.98 tot 3.39, gemixt bewijs, lage kwaliteit), terwijl IVF/ICSI resulteerden in een hogere kans op een levend geborene/doorgaande zwangerschap (OR 2.87, 95% BI 1.31 tot 6.29, gemixt bewijs, lage kwaliteit). Dit suggereert dat de als de kans op een levend geborene/doorgaande zwangerschap na een afwachtend beleid 15.9% is, de kans na OS, IUI, OS-IUI, IVF/ICSI respectievelijk 9.3%-29.6%, 11.8%-35.9%, 15.6%-39.0%, en 19.8%-54.3% zou zijn. De "surface under the cumulative ranking curve" (SUCRA) waarden voor een afwachtend beleid, OS, OS-IUI, IVF/ICSI waren respectievelijk 14.9%, 24.4%, 45.5%, 67.3%, 97.9%. Dit suggereert dat IVF/ICSI is de meest effectieve interventie als het gaat om doorgaande

zwangerschap of levend geborene, gevolgd door OS-IUI, IUI, OS en een afwachtend beleid.

In een subgroep analyse onderzochten wij paren met een duur van subfertiliteit van minder dan 2 jaar. In deze groep vonden wij geen bewijs van een voordeel van OS-IUI (OR 0.82, 95% BI 0.45 - 1.49, gemixt bewijs, lage kwaliteit) of IVF/ICSI (OR 1.05, 95% CI 0.46 - 2.43, indirect bewijs, lage kwaliteit) ten op zichten van een afwachtend beleid. In paren met een duur subfertiliteit van 2 jaar of meer, zowel OS-IUI (OR 2.84, 95% BI 1.25 - 6.43, gemixt bewijs, lage kwaliteit) als IVF/ICSI (OR 4.60, 95% BI 1.74 - 12.18, gemixt bewijs, lage kwaliteit) resulteerde in meer levend geborenen of doorgaande zwangerschappen.

Vergeleken met IUI, was de kans op een meerlingzwangerschap hoger bij OS-IUI en IVF/ICSI (OR 6.24, 95% BI 1.31 tot 29.69, gemixt bewijs; OR 5.89, 95% BI 1.18 tot 29.38, gemixt bewijs). Dit suggereert dat als de kans op een meerlingzwangerschap bij IUI 0.7% is, de kans na OS-IUI en IVF/ICSI respectievelijk 0.9%-17.7% en 0.8%-17.5% zou zijn. De SUCRA waarden voor een afwachtend beleid, OS, IUI, OS-IUI en IVF/ICSI waren respectievelijk 70.7%, 43.2%, 94.0%, 17.6% en 24.5%. Dit suggereert dat IUI en een afwachtend beleid het veiligste zijn met betrekking tot de kans op een meerlingzwangerschap, gevolgd door OS, IVF/ICSI en OS-IUI. Dit bewijs was van lage tot gemiddelde kwaliteit. Dit kwam met name door relatief weinig direct bewijs.

Geconcludeerd, bij paren met onverklaarde subfertiliteit leidt OS-IUI en IVF/ICSI tot meer levend geborenen of doorgaande zwangerschappen, maar ook tot een hoger aantal meerlingzwangerschappen.

In **hoofdstuk 3** geven wij de resultaten van een gerandomiseerde studie naar de effectiviteit van follikel stimuleren hormoon (FSH) vergeleken met clomifeen citraat (CC) bij paren met onverklaarde subfertiliteit die een behandeling met intra uteriene inseminatie (IUI) ondergaan met een maximum van vier cycli binnen zes maanden. De kern van deze studie was het protocol waarbij strikte cancelcriteria werden gehandhaafd om het risico op een meerlingrisico te verlagen. De cyclus werd gecanceld indien er meer dan drie dominante follikels groeiden. The primaire uitkomstmaat was het aantal doorgaande zwangerschappen en een belangrijke secundaire uitkomstmaat was het aantal meerlingzwangerschappen. 738 paren werden gerandomiseerd tussen FSH (n=369) en CC (n=369). In de FSH behandelarm ontstonden 113 doorgaande zwangerschappen (31%) en in de CC behandelarm ontstonden 97 doorgaande zwangerschappen (26%) (RR 1.16, 95% BI 0.93 - 1.47). In de FSH behandelarm ontstonden 5 meerlingzwangerschappen (1%) en in de CC behandelarm ontstonden 8 meerlingzwangerschappen (2%) (RR 0.63, 95% BI 0.21 - 1.89). Dit heeft geleid tot de conclusie dat een protocol waarbij strikte cancelcriteria worden gehandhaafd een succesvolle oplossing is om IUI met ovariële stimulatie uit te voeren, zonder vrouwen bloot te stellen aan een hoog meerlingrisico.

In **hoofdstuk 4** beschrijven wij de resultaten van een netwerk meta-analyse waarin de effectiviteit en veiligheid van IUI met clomifeen citraat, Letrozole of gonadotrofinen met elkaar wordt vergeleken en met IUI in een natuurlijke cyclus in paren met onverklaarde subfertiliteit. Wij includeerden 26 gerandomiseerde studies (5 316 vrouwen). Hoewel IUI met gonadotrofinen het hoogste uit de ranking kwam wat betreft doorgaande zwangerschap/levend geborene, vrouwen die deze behandeling ondergingen liepen ook een risico op een meerlingzwangerschap met een hoog complicatierisico. Een IUI protocol waarbij strikte cancelcriteria worden gehanteerd leidt tot een acceptabel meerlingrisico, zonder dat de effectiviteit hieronder leidt. Indien een protocol met strikte cancelcriteria wordt gevolgd, leidt het gebruik van gonadotrofinen tot meer doorgaande zwangerschappen/levend geborene in vergelijking met clomifeen citraat. Wij adviseren daarom om bij paren met onverklaarde subfertiliteit IUI uit te voeren met strikte cancelcriteria en te stimuleren met gonadotrofinen. Hierbij houden wij geen rekening met kosten van de verschillende medicijnen voor ovariële stimulatie.

In **hoofdstuk 5** hebben wij beoogd om individuele selectie markers vast te stellen om paren te identificeren die baat hebben bij specifiek gonadotrofinen of clomifeen citraat bij IUI vanwege onverklaarde subfertiliteit. We hebben hiervoor een secundaire analyse van de SUPER studie uit hoofdstuk 3 uitgevoerd. As potentiele selectie markers hebben wij gekeken naar vrouwelijke leeftijd, BMI, duur van subfertiliteit, of er sprake was van primaire of secundaire subfertiliteit en semen kwaliteit. Voor elke factor hebben wij een logistisch regressiemodel opgesteld. Geen van de voorgestelde selectie markers bleek geassocieerd te zijn met hogere doorgaande zwangerschap cijfers bij gonadotrofinen in vergelijking met clomifeen citraat. Op basis van deze data kunnen wij dus geen specifieke subgroepen van vrouwen identificeren die baat hebben bij gonadotrofinen boven clomifeen citraat.

In **hoofdstuk 6** evalueerden wij de impact van gonadotrofinen en clomifeen citraat op de endometrium dikte en de associatie tussen endometrium dikte en de kans op doorgaande zwangerschap. We hebben hiervoor een secundaire analyse van de SUPER studie uit hoofdstuk 3 uitgevoerd. We hebben het verschil in endometriumdikte tussen vrouwen die waren toegewezen voor IUI met gonadotrofinen en vrouwen die waren toegewezen voor IUI met clomifeen citraat berekend met een lineair gemixt model. De associatie tussen endometrium dikte en de kans op doorgaande zwangerschap hebben we uitgerekend met een logisch regressie model, waarbij we gecorrigeerd hebben wij bekende voorspellers voor zwangerschap. 666 paren ondergingen 1968 IUI cycli. De gemiddelde endometriumdikte was 8.9 mm (SD 2.1) bij vrouwen die toegewezen waren aan IUI met gonadotrofinen en 7.5 mm (SD 2.1) in vrouwen die waren toegewezen aan IUI met clomifeen citraat (gemiddeld verschil van 1.4mm, 95% BI 1.1 – 1.7). De gemiddelde endometriumdikte was 8.4 mm (SD 2.2) bij vrouwen met een doorgaande zwangerschap en 8.2 mm (SD 2.2) bij vrouwen die niet zwanger zijn geworden (gemiddeld verschil 0.29mm, 95% BI -0.10 – 0.68). We vonden geen duidelijke associatie tussen endometriumdikte en doorgaande zwangerschap (odds ratio 1.07

mm per 1 mm, 95% BI 0.99 – 1.15). Bij vrouwen die IUI ondergaan in het kader van onverklaarde subfertiliteit is een dun endometrium geen reden om te wisselen van medicijn voor ovariële stimulatie.

In **hoofdstuk 7** beschrijven wij de resultaten van een kosteneffectiviteit analyse van de SUPER studie uit hoofdstuk 3. Deze studie is uitgevoerd vanuit een gezondheidszorg perspectief. We hebben directe medische kosten vergeleken van gonadotrofinen en clomifeen citraat bij paren met onverklaarde subfertiliteit die IUI ondergaan met strikte cancelcriteria. Directe medische kosten bestonden uit kosten voor medicatie, cyclus monitoring, inseminatie en indien van toepassing zwangerschapscontroles. We berekenden gemiddelde kosten per paar en de incrementele kosten effectiviteit ratio (ICER) tussen gonadotrofinen en clomifeen citraat. De gemiddelde kosten per paar waren €1 534 voor gonadotrofinen en €1 067 voor clomifeen citraat (gemiddeld verschil €468 (95% BI, €464 - €472)). De incrementele kosten effectiviteit ratio was €21 804 (95% BI, €11 628 - €31 980) per extra doorgaande zwangerschap bij IUI met gonadotrofinen. In een systeem waarbij voortplantingszorg uit private middelen betaald dient te worden is het belangrijk deze data aan de patiënten de communiceren. In een systeem waarbij voortplantingszorg wordt betaald uit publieke middelen, het is de vraag of €21 804 moet worden uitgegeven aan een extra doorgaande zwangerschap. Dit is een vraag die aan de kaak moeten worden gesteld bij zorginkopers.

REFERENTIE

Evers JL. Female subfertility. *Lancet* 2002;360:151-159.

A.

List of coauthors and affiliations

Author contributions

PhD portfolio

List of publications

Dankwoord

About the author

LIST OF COAUTHORS AND AFFILIATIONS

Bhattacharya	**S**	Cardiff University School of Medicine, Heath Park, Cardiff, United Kingdom Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom
Boks¥	**DES**	Department of Obstetrics and Gynaecology, Spaarne Hospital, Hoofddorp, the Netherlands
Bossuyt	**PMM**	Department of Epidemiology, Biostatistics & Bioinformatics, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
Broekmans¥	**FJM**	Centre for Reproductive Medicine, University Medical Centre Utrecht, Utrecht, the Netherlands
Cohlen¥	**BJ**	Isala Hospital. Department of Obstetrics and Gynaecology, Zwolle, the Netherlands
de Bruin¥	**JP**	Jeroen Bosch Hospital, Department of Obstetrics and Gynaecology, Den Bosch, the Netherlands
Eijkemans	**MJC**	Department of Biostatistics and Research Support, Julius Centre, University Medical Centre Utrecht, Utrecht, the Netherlands
Fleischer¥	**K**	Centre for Reproductive Medicine, Radboud University Medical Centre, Nijmegen, the Netherlands
Gianotten¥	**J**	Department of Obstetrics and Gynaecology, Spaarne Hospital, Haarlem, the Netherlands
Janssen¥	**CAH**	Department of Obstetrics and Gynaecology Groene Hart Hospital, Gouda, the Netherlands
Klijn¥	**N**	Department of Obstetrics and Gynaecology, Leiden University Medical Centre, Leiden, the Netherlands
Koks¥	**C**	Department of Obstetrics and Gynaecology, Máxima Medical Centre, Veldhoven, the Netherlands
Mochtar¥	**MH**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
Mol¥	**BW**	Department of Obstetrics and Gynaecology, Monash University, Clayton, Australia
Pelinck¥	**MJ**	Department of Obstetrics and Gynaecology, Scheper Hospital, Emmen, the Netherlands
Perquin¥	**DAM**	Department of Obstetrics and Gynaecology Medical Centre Leeuwarden, Leeuwarden, the Netherlands
Repping¥	**S**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
Rijn -van Weert¥	**JM**	Department of Obstetrics and Gynaecology, NoordWest Groep Alkmaar, Alkmaar, the Netherlands

Sluijmer¥	**A**	Department of Obstetrics and Gynaecology Wilhelmina hospital, Assen, the Netherlands
Tjon-Kon-Fat	**RI**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
Traas¥	**MAF**	Department of Obstetrics and Gynaecology Gelre Hospital, Apeldoorn, the Netherlands
Twisk¥	**M**	Department of Obstetrics and Gynaecology, Amstelland hospital, Amstelveen, the Netherlands
van der Ham¥	**DP**	Department of Obstetrics and Gynaecology, Martini hospital, Groningen, the Netherlands
van der Veen¥	**F**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
van Disseldorp¥	**J**	Department of Obstetrics and Gynaecology, St. Antonius hospital Nieuwegein, Nieuwegein, the Netherlands
van Eekelen	**R**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
van Heteren¥	**CF**	Department of Obstetrics and Gynaecology, Canisius Wilhelmina Hospital, Nijmegen, the Netherlands
van Hooff¥	**MHA**	Department of Obstetrics and Gynaecology Sint Franciscus Gasthuis, Rotterdam, the Netherlands
van Schaik	**RHN**	Department of Clinical Chemistry, Erasmus MC, Rotterdam, the Netherlands
van Wely¥	**M**	Centre for Reproductive Medicine, Amsterdam UMC, location Academic Medical Centre, Amsterdam, the Netherlands
Verberg¥	**MFG**	Fertility Clinic Twente, Hengelo, the Netherlands
Verhoeve¥	**HR**	Department of Obstetrics and Gynaecology, OLVG location Oost, Amsterdam, the Netherlands
Visser¥	**J**	Department of Obstetrics and Gynaecology Amphia, Breda, the Netherlands
Wang	**R**	Department of Obstetrics and Gynaecology, Monash University, Clayton, Australia

¥ Member of the SUPER study group

AUTHOR CONTRIBUTIONS

Chapter 2

RW, RIT, MJCE, PMMB, FvdV, SB, BWM and MvW contributed to the study design and the original protocol. RW, ND, RIT and MvW collected the data. RW and MvW analysed the data. All authors interpreted the data. RW wrote the first draft. All authors revised the manuscript critically for important intellectual content and approved the final version.

Chapter 3

ND was responsible for the overall logistical aspects of the trial and drafted the paper. MM, FvdV and MW designed the trial and were responsible for the development of the protocol. All authors contributed to the protocol and approved the final version of the article.

Chapter 4

ND was responsible for the overall logistical aspects of the study and drafted the paper. ND, RW, BWM, MM, FvdV and MW designed the study. RW was responsible for the statistical analysis. All authors contributed and approved the final version of the paper

Chapter 5

ND was responsible for the overall logistical aspects of the study and drafted the paper. MM, FvdV and MW designed the study. RvE was responsible for the statistical analysis. All authors contributed and approved the final version of the paper.

Chapter 6

ND was responsible for the overall logistical aspects of the study and drafted the paper. MM, FvdV and MW designed the study. RvE was responsible for the statistical analysis. All authors contributed and approved the final version of the paper.

Chapter 7

ND was responsible for the overall logistical aspects of the study and drafted the paper. MM, FvdV and MvW designed the study. ND, MvW, SR, DvdH, NK, CJ, JRvW, MTw, MTr, MP, DP, DB, AS, BWM, FvdV and MM contributed and approved the final version of the paper.

PHD PORTFOLIO

Courses

2015	BROK: legislation and organization for clinical researchers, 1.0 ECTS
2016	Practical Biostatistics, 1.1 ECTS
2016	Research data management, 0.7 ECTS
2016	Clinical Epidemiology: Randomized Clinical Trials, 0.6 ECTS
2017	Scientific writing in English for publication, 1.5 ECTS
2018	Computing in R, 0.4 ECTS

Seminars

2015 – 2019	Weekly department seminars, department of Obstetrics and Gynaecology, AUMC, location AMC
2015	Symposium on prediction models, Centre for Reproductive Medicine, Amsterdam, The Netherlands
2016	Consortium training day, UMCG, Groningen
2016	8 years TRUST symposium, Centre for Reproductive Medicine, Amsterdam, The Netherlands
2019	Symposium update legislation and organization for clinical researchers (re-registration BROK)

Oral presentations

2017	Ovarian stimulation in intrauterine insemination
	European Society of Human Reproduction and Embryology (ESHRE) annual meeting, July 2-5 2017, Geneva, Switzerland
2017	Ovariële stimulatie bij intrauteriene inseminatie
	Invited speaker, 52th Gynaecongres, Nederlandse vereniging voor Obstetrie en Gynaecologie (NVOG), Amersfoort, The Netherlands
2018	A cost-effectiveness analysis of follicle stimulating hormone compared with clomiphene citrate in intra uterine insemination
	34th European Society of Human Reproduction and Embryology (ESHRE) annual meeting, July 1-4 2018, Barcelona, Spain 2018

2018	What is in unexplained subfertility the impact of gonadotrophins or clomiphene citrate on endometrial thickness, a secondary analysis of a trial on intrauterine insemination 34th European Society of Human Reproduction and Embryology (ESHRE) annual meeting, July 1-4 2018, Barcelona, Spain 2018
2018	Interventions for unexplained infertility: a systematic review and network meta-analysis Dutch society of Fertility Doctors meeting
2019	Ovariële stimulatie bij intrauteriene inseminatie Fertility Club North Holland

Poster presentations

2013	The validity of the variable "NICU admission" as an outcome measure for neonatal morbidity: a retrospective cohort study 33rd Society for Maternal-Fetal Medicine (SMFM) annual meeting, February 11-16 2013, San Francisco, California
2016	Recombinant luteinizing hormone (rLH) and recombinant follicle stimulating hormone (rFSH) for ovarian stimulation in IVF ICSI cycles European Society of Human Reproduction and Embryology (ESHRE) annual meeting, July 3-6 2016, Helsinki, Finland

Conferences

2013	33rd Society for Maternal-Fetal Medicine (SMFM) annual meeting February 11-16 2013, San Francisco, California
2015	48th Gynaecongres, Nederlandse vereniging voor Obstetrie en Gynaecologie (NVOG), Arnhem, The Netherlands
2016	50th Gynaecongres, Nederlandse vereniging voor Obstetrie en Gynaecologie (NVOG), Amersfoort, The Netherlands
2017	52th Gynaecongres, Nederlandse vereniging voor Obstetrie en Gynaecologie (NVOG), Amersfoort, The Netherlands
2018	34th European Society of Human Reproduction and Embryology (ESHRE) annual meeting, July 1-4 2018, Barcelona, Spain

LIST OF PUBLICATIONS

Litjens RJNTM, Oude Rengerink K, **Danhof NA**, Kruitwagen RFPM, Mol BWJ. Does recruitment for multicenter clinical trials improve dissemination and timely implementation of their results? A survey from the Netherlands. Clinical Trials, 2013;10(6):915-23

Wiegerinck MMJ*, **Danhof NA***, van Kaam AHL, Tamminga P, Mol BWJ. The validity of the variable "NICU admission" as an outcome measure for neonatal morbidity. *both authors contributed equally to this study, Acta Obstet Gynaecol Scan. 2014 Jun;93(6):603-9

Danhof NA, Kamphuis EI, Mol BWJ. Loop electrosurgical Excision Procedure and Risk of Preterm Birth, a letter to the editor. Obstet Gynecol. 2014 Jul;124(1):163

Danhof NA, Kamphuis EI, Limpens J, Lonkhuijzen LRCW, Pajkrt E, Mol BWJ. The risk of preterm birth of treated versus untreated cervical intraepithelial neoplasia (CIN): a systematic review and meta-analysis. European Journal of Obstetrics and Gynecology and Reproductive Biology 2015; 188:24 – 33

Mochtar MH, **Danhof NA**, Ayeleke RO, Van der Veen F, van Wely M. Recombinant luteinizing hormone (rLH) and recombinant follicle stimulating hormone (rFSH) for ovarian stimulation in IVF/ICSI cycles. Cochrane Database Syst Rev. 2017 May 24;5:CD005070.

Danhof NA, van Wely M, Koks CAM, Gianotten J, de Bruin JP, Cohlen BJ et al. The SUPER study: protocol for a randomised controlled trial comparing follicle-stimulating hormone and clomiphene citrate for ovarian stimulation in intrauterine insemination. BMJ Open. 2017 May 25;7(5):e015680

Kamphuis EI, Naber SK, **Danhof NA**, Habbema JDF, de Groot CJM, Mol BWJ Effect of Cervical Cancer Screening Programs on Preterm Birth: A Decision and Cost-Effectiveness Analysis. Obstet Gynecol. 2017 Dec;130(6):1207-1217

Danhof NA, van Wely M, Repping S, Koks C, Verhoeve HR, de Bruin JP, Verberg MFG, van Hooff MHA, Cohlen BJ, van Heteren CF, Fleischer K, Gianotten J, van Disseldorp J ,Visser J, Broekmans FJM, Mol BWJ, van der Veen F, Mochtar MH, for the SUPER study group. Follicle stimulating hormone (FSH) versus Clomiphene Citrate (CC) in intrauterine insemination for unexplained subfertility: A randomized controlled trial. Human reproduction 2018;33(10):1866-74

Danhof NA, van Eekelen R, Repping S, Mol BWJ, van der Veen F, van Wely M, Mochtar MH, for the SUPER study group. FSH or CC in IUI with ovarian stimulation for unexplained subfertility: a role for treatment selection markers? Reprod Biomed Online. 2019 Jun;38(6):938-942

Wang R, **Danhof NA**, Tjon-Kon-Fat RI, Eijkemans MJC, Bossuyt PMM, Mochtar MH, van der Veen F, Bhattacharya S, Mol BWJ, van Wely M. Interventions for unexplained infertility: a systematic review and network meta-analysis. Cochrane Database Systematic Reviews. 2019;9:CD012692

Danhof NA, Wang R, van Wely M, van der Veen F, Mol BWJ, Mochtar MH. Intrauterine insemination for unexplained infertility – a network meta-analysis. Human Reproductive Update. 2020;26(1):1-15.

Danhof NA, van Eekelen R, Repping S, Mol BWJ, van der Veen F, van Wely M, Mochtar MH, for the SUPER study group. Endometrial thickness as a biomarker for ongoing pregnancy in intrauterine insemination for unexplained subfertility: a secondary analysis. Human Reproduction Open. In press.

Danhof NA, van Wely M, Repping S, van der Ham DP, Klijn N, Janssen CAH, Rijn-van Weert JM, Twisk M, Traas MAF, Pelinck MJ, Perquin DAM, Boks DES, Sluijmer A, Mol BWJ, van der Veen F, Mochtar MH, for the SUPER study group. Gonadotrophins or clomiphene citrate in couples with unexplained infertility undergoing intrauterine insemination: a cost-effectiveness analysis. Reproductive Biomedicine Online. 2019;19: pii: S1472-6483,30740-0

DANKWOORD

Mijn proefschrift is af! Dit kon ik natuurlijk niet alleen. Jarenlang heb ik er samen met anderen met plezier aan gewerkt. In dit dankwoord richt ik mij graag tot een aantal personen in het bijzonder.

Ik wil alle vrouwen bedanken die hebben meegewerkt aan het onderzoek in dit proefschrift.

Beste Fulco, bedankt voor je grenzeloze inzet om mij verder te helpen in het onderzoek. Ik heb veel van je geleerd, onder andere dat ik nooit met minder dan het hoogst haalbare genoegen moet nemen. Dankzij je eerlijkheid heb ik geleerd om kritisch te zijn.

Beste Sjoerd, bedankt voor onze inspirerende gesprekken en brainstormsessies. Je stimuleerde mij om alles net van een andere kant te bekijken. Ik heb fijn met je samengewerkt.

Beste Monique, bedankt voor de kans om bij het CVV te komen werken. Met jou heb ik de eerste stappen als arts en als onderzoeker gezet. Volgens mij zijn wij in een korte tijd een goed team geworden. Moet ik nu echt afscheid nemen? De deur stond altijd open. Zonder jou was het mij niet gelukt om nu hier te staan.

Beste Madelon, gelukkig wilde ook jij copromotor worden! Wanneer ik mij door data en statistiek worstelde hielp jij mij uit de brand. Met al je ervaring was je ook een fijn luisterend oor over alles met promoveren te maken heeft.

Geachte leden van de promotiecommissie: prof. dr. M. Goddijn, prof. dr. V. Mijatovic, prof. dr. C.B. Lambalk, prof. dr. M.J.C. Eijkemans, prof. dr. J.S.E. Laven, graag wil ik u bedanken voor de bereidheid mijn proefschrift kritisch te beoordelen.

Graag wil ik ook alle medeauteurs bedanken voor hun waardevolle bijdrage aan dit proefschrift.

Dank aan alle (research-)medewerkers van het Consortium 2.0 voor de ondersteuning en fijne samenwerking bij het volbrengen van de SUPER studie.

Dan ‘mijn double R’, Rik en Rui! Bijna al mijn projecten heb ik samen met jullie gedaan. Rik, onze proefschriften liggen qua onderwerp dicht bij elkaar, terwijl wij allebei een totaal andere achtergrond hebben. Dank je dat je altijd voor mij klaar stond en mij verder hebt geholpen. En natuurlijk bedankt voor de gezelligheid op H4! Rui, you are located in Australia and I was located in Amsterdam, which made it possible to work clock round. When we worked on a manuscript I send it to you when my day ended and the next day I found your new draft in my mailbox. Thank you for helping me and for our cozy skype chats.

Beste Ben Willem, bedankt dat je mij als student de kans hebt gegeven om te beginnen als onderzoeker. Het was een hele welkome en leerzame start voor het werk in dit proefschrift.

Lieve collega's van het CVV, lieve gynaecologen, fertiliteitsartsen, verpleegkundigen, doktersassistente en iedereen van de back-office, mijn allereerste baan begon bij jullie en wat heb ik een goede tijd gehad. Bedankt dat ik bij de CVV-familie mocht horen. Jullie waren meer dan collega's. Lieve roze muisjes, lieve Miriam, Judith, Odine en Merel, ik heb vriendinnen voor het leven gemaakt.

Lieve (arts-)onderzoekers van H4. Ik wil jullie allemaal bedanken voor de goede tijd.

Beste Michelle, bedankt dat ik de SUPER studie van je kon overnemen. Je hebt de studie strak opgezet, waardoor de inclusies binnen stroomden.

Aan alle collega's van OLVG Oost, bedankt voor de leerzame en bovenal gezellige tijd. Jullie hebben mij de eerste kneepjes van het gynaecologen vak geleerd.

Aan alle collega's van cluster LUMC, bedankt dat ik de kans heb gekregen om hier mijn opleiding te starten. Ik voel mij heel welkom.

Aan alle collega's van het Groene Hart Ziekenhuis, ik ben in een warm bad terecht gekomen. Bedankt voor de fijne start van de opleiding.

Lieve Jasper Kroese, bedankt voor alle tijd en het geduld die je in het design van dit proefschrift hebt gestoken.

Lieve vriendinnen, en dan in het bijzonder in willekeurige volgorde: Yara, Henriette, Willemijn, Leontine, Sophie H, Nicky, Natalja, Catherine, Isabelle, Emma, Ilse, Amanda, Valery, Veronique, Djoeke, Priyanka, Sophie D, Philippine, Biba, Hanke, wat ben ik blij en trots met jullie in mijn buurt. Ik hoop dat we de rest van onze levens met elkaar kunnen delen!

Lieve paranimfen (maar liefst drie!), gelukkig staan jullie naast mij vandaag. Jullie zijn voor mij onmisbaar geweest om dit proefschrift te schrijven. Lieve Mir, buddy bij het CVV, met jouw steun bij de start bij het CVV, het onderzoek en de start in het OLVG ben jij stiekem een mentor voor mij geworden en je bent er altijd als ik het nodig heb. Je bent een hele lieve attente vriendin. Lieve Juud, wij gingen de afgelopen jaren ongeveer gelijk op, waardoor we veel aan elkaar gehad hebben. Jij hebt het woord "collega-vriendin" geïntroduceerd, maar dat is nu wel echt vriendin geworden. Lieve Eef, wij hebben bijna een jaar naast elkaar aan tafel gezeten en toen hebben we veel life-events gedeeld, trouwen, huis kopen en baby's krijgen. Dat levert een bijzondere band op.

Lieve schoonfamilie, bedankt voor alle gezelligheid, etentjes, vakanties en dat jullie altijd klaar staan.

Lieve papa, mama, Job en Sanne, bedankt voor de basis in mijn leven. Ik heb veel aan jullie te danken. Papa en mama, jullie zijn er altijd als ik het nodig heb en even vastloop. Bedankt voor al het meedenken, de wijze lessen en alle liefde. Ik hou van jullie.

Lieve Jas en Filip, jullie zijn voor mij waar het om draait in het leven. Lieve Filip, voordat jij er was kon ik mij niet voorstellen hoe bijzonder het is om een kindje te krijgen. Ik kan geen genoeg van je krijgen. Laten we nog heel veel boekjes lezen, spelen en knuffelen. Lieve Jas, ik hou nog steeds elke dag meer van je. Wij hebben zoveel plezier samen. Samen kunnen wij de hele wereld aan. Jij bent de beste partner en de beste vader van Filip die ik mij kan wensen. Over een paar maanden begint ons nieuwe grote avontuur. Ik kan niet wachten!

ABOUT THE AUTHOR

Noor Danhof was born on the 11th of September in 1988 in Leiden, the Netherlands. In 2006 she graduated from secondary school at the Stedelijk Gymnasium, after which she took a year off to travel in South-East Asia. She moved to Amsterdam in 2007 where she started studying Human Movement Sciences at the Vrije Universiteit in Amsterdam. One year later she was accepted to start with medical school at the University of Amsterdam. During her study she followed an internship on the department of Obstetrics and Gynaecology at Kings College Londen, UK. In February 2015 she graduated from medical school and started working at the Centre for Reproductive Medicine at the Amsterdam University Medical Centre location AMC as a fertility doctor and a PhD student under the supervision of Prof. dr. F. van der Veen, Prof. dr. S. Repping, Dr. M.H. Mochtar and Dr. M. van Wely. In 2018 she started working as a resident not in training on the department of Obstetrics and Gynaecology at the OLVG Oost in Amsterdam. She currently works as a resident in training, specializing in Obstetrics and Gynaecology, at the Groene Hart Hospital in Gouda under the supervision of Dr. C. van Meir and Dr. M. Sueters.

She lives together with her partner Jasper and their son Filip in Leiden.