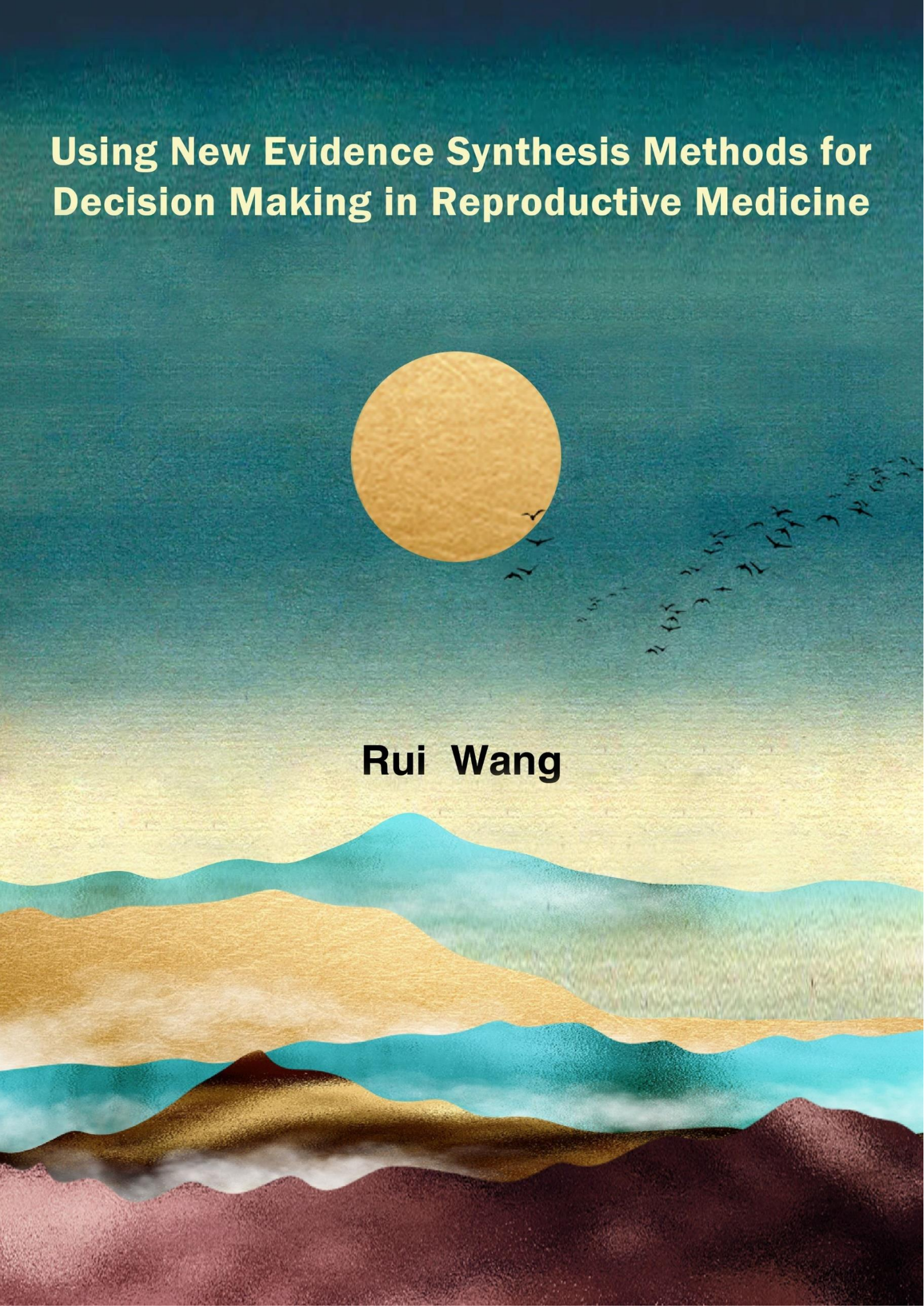
Using New Evidence Synthesis Methods for Decision Making in Reproductive Medicine



Rui Wang

Using New Evidence Synthesis Methods for Decision Making in Reproductive Medicine

Rui Wang

Robinson Research Institute

Adelaide Medical School

The University of Adelaide

September 2019

TABLE OF CONTENTS

LIST OF INCLUDED PUBLICATIONS

1. **Wang R**, Seidler AL, Askie L, Norman RJ, Bhattacharya S, van Wely M, Mol BWJ. Network meta-analyses in reproductive medicine: challenges and opportunities. 2019. Unsubmitted (Chapter 2)
2. **Wang R**, Kim BV, van Wely M, Johnson NP, Costello MF, Zhang H, Ng EH, Legro RS, Bhattacharya S, Norman RJ, Mol BW. Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis. *BMJ*. 2017;356:j138. (Chapter 3)
3. **Wang R**, Jones ER, Costello MF, Bhattacharya S, Legro RS, Ng E, Johnson NP, Norman RJ, van Wely M, Mol BWJ. Ovulation induction for women with polycystic ovary syndrome treated unsuccessfully with clomiphene citrate: a systematic review and network meta-analysis. 2019. Unsubmitted. (Chapter 4)
4. **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 of Systematic Reviews*. 2019, 9. Art. No.: CD012692. DOI: 10.1002/14651858.CD012692.pub2. (Chapter 5)
5. **Wang R**, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, Mijatovic V, Mol BW. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. *Ultrasound Obstet Gynecol*. 2019; 54(2):172-181. (Chapter 6)
6. **Wang R**, Mol BWJ. The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria? *Human Reproduction*. 2017;32(2):261-264. (Chapter 7)
7. **Wang R**, Li W, Bordewijk EM, Legro RS, Zhang H, Wu X, Gao J, Morin-Papunen L, Homburg R, König TE, Moll E, Kar S, Huang W, Johnson NP, Amer SA, Vegetti W, Palomba S, Falbo A, Özmen Ü, Nazik H, Williams CD, Federica G, Lord J, Sahin Y, Bhattacharya S, Norman RJ, van Wely M, Mol BW, Reproductive Medicine Network, the International Ovulation Induction IPDMA Collaboration. First-line ovulation induction for polycystic ovary syndrome: an individual participant data meta-analysis. *Human Reproduction Update*. 2019. Accepted. (Chapter 8)

ABSTRACT

Pairwise meta-analyses with aggregate data are the current standard method for evidence synthesis in evidence-based medicine. Network meta-analyses and individual participant data (IPD) meta-analyses are relatively new evidence synthesis methods for decision making in health care. Network meta-analyses involve simultaneous comparisons of multiple interventions and IPD meta-analyses provide insights into personalised medicine. These methods have the potential to overcome the drawbacks of conventional evidence synthesis methods and optimise the available evidence and therefore are promising for decision making in reproductive medicine. In this thesis, we applied network meta-analyses to answer a series of clinical questions in reproductive medicine, including the first-line and second-line treatment strategies for polycystic ovary syndrome, clinical managements of unexplained infertility and the use of different contrast media during tubal patency testing to improve fertility outcomes. In addition, we used IPD meta-analysis to identify personalised first-line treatment strategies for polycystic ovary syndrome to improve fertility outcomes. These results can be used to upgrade the current evidence base in reproductive medicine and provide a robust basis for clinical guideline development and directions for future research.

THESIS DECLARATIONS

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

PhD Candidate: Rui Wang

Signature:

Date: 12/9/2019

ACKNOWLEDGEMENT

First, I would like to thank my two supervisors Professor Ben Willem Mol and Professor Robert Norman for their guidance during my PhD journey.

Dear Ben Willem, thank you for guiding me to see the bigger picture in clinical research. It was a great challenge for me at the very beginning as I work according to "The devil is in the detail" principle. I spent a considerable amount of time at the start of my PhD in details and only gradually did I realise I have more important things to focus upon. Also thank you for introducing me to numerous amazing collaborators. I absolutely enjoyed the experience working with different people during my PhD.

Dear Rob, thank you for introducing me to Adelaide, providing me a great opportunity to work with extraordinary people around me. Thank you for being a fabulous mentor, inviting me to the cricket game at Adelaide Oval, guiding me to understand the Australian culture and supporting me during my tough PhD days. You inspired me to make the complicated methodology more assessible to clinicians and patients and to keep reminding myself for whom I am doing research.

I would like to thank all my co-authors for their help with my thesis, especially Dr Madelon van Wely from UMC Amsterdam and Professor Siladitya Bhattacharya from The University of Aberdeen. Madelon, thank you for all your remote methodological supports and for inviting me to visit Amsterdam and to meet with the fabulous "Golden Sperm" team in the AMC. It has been a great pleasure to work with you and I'm looking forward to all the future collaborations with you. Bhatty, thank you for being a great mentor in guiding me in the right direction, and for all the amazing dinners at ESHRE annual meetings. I wish you all the best in your new position in Aberdeen.

I would like to thank Professors Richard Legro, Neil Johnson, Heping Zhang, Ernest Ng, Velja Mijatovic and Fulco van der Veen, Dr Michael Costello and the members of the

Reproductive Medicine Network and the International Ovulation Induction IPDMA Collaboration for their support.

I would like to thank Dr Sean O'Leary for being a great mentor and friend in Adelaide. Thanks for your company in sharing the happiness and anxieties.

I would like to thank Noor Danhof, Nienke van Welie, Esmée M Bordewijk and Dr Raïssa Tjon-Kon-Fat from the Netherlands, for all the early morning Skype meetings and for all the help in my systematic reviews. You are fabulous colleagues and I wish you the best in completing your PhD and/or working in your new positions.

I would like to thank Bobae Kim and Esther Jones from the University of Adelaide for the help with my systematic reviews.

A big thank to all my friends on Level 3 Norwich Centre at Robinson Research Institute, Dr Jesia Berry, Dr Suzette Coat, Dr Anna Kontos, Kelly Harper, Tara Crawford and Amy Grace. Thank you for your company when I was in Adelaide and thanks for all the Friday afternoon drinks and chats. Thank you for the farewell dinner and I will wear the "Heaps Good" hoodie when submitting my thesis!

I would also like to thank the colleagues from Department of Obstetrics and Gynaecology, Monash University, especially Dr Wentao Li and Hildegard Mostmans for their support during my visit in Monash University. I'm glad that we will continue working closely for at least several more years!

Thank my cousin Jing for her support to my overseas study experience. Thanks for preparing such an excellent thesis cover. You are an extraordinary graphic designer!

Finally, I would like to thank my parents for their support during my PhD journey. I enjoy all the family time with you, cooking, dining, travelling and more!

Chapter 1

General introduction and outline of the thesis

GENERAL INTRODUCTION

Infertility, defined as the failure to achieve pregnancy after 12 months of unprotected intercourse, affects up to 186 million couples worldwide (Inhorn and Patrizio, 2015; Mascarenhas et al., 2012). Infertility and involuntary childlessness are significant life stressors that have serious consequences for couples' quality of life and emotional wellbeing, their families and society as a whole (Rockliff et al., 2014).

Reproductive medicine is a relatively young discipline in medicine, aiming to help the above-mentioned couples fulfil their child wish by providing effective and safe treatment. Such process of clinical decision-making in reproductive medicine is not different from that in other medical disciplines - applying evidence-based medicine principles to find the best available treatment(s).

In evidence-based medicine, randomised controlled trials (RCTs) are performed to compare the effectiveness as well as the safety of two or more treatment options. When multiple primary RCTs are available, systematic reviews and conventional meta-analyses are conducted to summarise the available evidence to help clinical decision making (Higgins and Green, 2011). In a systematic review, a comprehensive electronic search is performed to identify all the RCTs that could answer the research question. Next, RCTs fulfilling the inclusion are critically assessed. The methodological quality of the included RCTs is then evaluated by using standardised scale system such Cochrane Collaboration's tool (Higgins and Green, 2011). Finally, statistical methods are used to summarise the overall findings, the process of which is named meta-analysis (Higgins and Green, 2011).

Traditional pairwise aggregated data meta-analysis only allows the comparison of two different interventions. When several treatment options are available for a clinical condition or a disease, multiple pairwise meta-analyses can be performed. However, clinical decision making based on these multiple meta-analyses does not seem clear, as a single meta-analysis of two interventions does not reflect the whole picture. When multiple treatment options

have been compared for a specific disease and the same outcomes, network meta-analysis, also known as multiple treatment comparison meta-analysis, is able to make evidence of multiple interventions that is both easy to visualise and able to be interpreted in a wider picture (Lu and Ades, 2004; Mills et al., 2013; Song et al., 2009). Network meta-analysis allows the comparison of multiple interventions simultaneously for a specific clinical condition by using both direct and indirect evidence, and to generate rankings for both the effectiveness and safety of available interventions to guide clinical decision making (Jansen et al., 2011; Salanti, 2012).

Another key limitation of traditional pairwise meta-analysis based on aggregated data is that the effectiveness of intervention is evaluated at a group level rather than at an individual level and therefore the recommendation based on these meta-analyses follows "one-size-fits-all" principle. In clinical practice, it is important to identify which individuals benefit most from a particular treatment so that clinicians can provide personalised care. However, such a question on subgroup effects, although clinically important, is usually impossible to be answered in primary RCTs due to the underpowered nature of subgroup analysis (Riley et al., 2010). Neither can it be solved in meta-analyses due to ecological bias resulting from the ignorance of within-study interactions or limited data availability resulting from heterogeneous reporting of subgroup data in the primary RCTs (Riley et al., 2010). Individual participant data (IPD) meta-analysis has the potential to solve the above-mentioned clinical dilemma, as the data synthesis process is achieved by using IPD from RCTs. The inclusion and exclusion criteria can be standardised and the statistical analysis can be performed consistently across all the included RCTs in IPD meta-analysis (Riley et al., 2010). Moreover, it is more flexible is to choose relevant endpoints, including time-to-event outcomes, and to investigate subgroups effects or treatment-covariate interactions (Broeze et al., 2010; Thompson and Higgins, 2005).

Network meta-analysis and IPD analysis are novel evidence synthesis methods in clinical decision making. They have the potential to overcome the drawbacks of traditional pairwise meta-analysis based on aggregated data and extend the dimensions of evidence synthesis processes by incorporating multiple treatment comparisons and facilitating personalised treatment choices. Therefore, these are promising methods in decision making in evidence-based reproductive medicine.

In this thesis, we first introduce the basic principle of network meta-analysis (Chapter 2) and then apply this method to solve four clinical questions (Chapter 3-6), including the first-line, second-line ovulation induction for WHO group II anovulation / polycystic ovary syndrome (PCOS), clinical management of unexplained infertility and choices of different contrast media during tubal patency tests. Next, we address the need of IPD meta-analysis to guide personalised treatment choice for PCOS (Chapter 7) and use IPD meta-analysis to address the clinical question on personalised ovulation induction strategy for PCOS (Chapter 8). Finally, we provide discussions, implications and directions for future research.

OUTLINE OF THE THESIS

In **chapter 2**, we introduced network meta-analyses and the underlining key assumption. We provided an overview of published network meta-analyses in reproductive medicine and identified the research gaps and opportunities.

In **chapter 3**, we performed a network meta-analysis to compare different ovulation induction strategies as the first-line treatment for women with WHO group II anovulation and summarised the evidence base for clinical practice.

In **chapter 4**, we performed a network meta-analysis to compare different ovulation induction strategies as the second-line treatment for women with PCOS who are treated unsuccessfully with clomiphene citrate and provided guidance for clinical practice.

In **chapter 5**, we performed a network meta-analysis to compare different treatment strategies, including expectant management, ovarian stimulation (OS), intrauterine insemination (IUI), OS-IUI, and IVF or ICSI for couples with unexplained infertility, and identified the research gap on unexplained infertility.

In **chapter 6**, we performed a network meta-analysis to compare the use of different contrast media during tubal testing on fertility outcome and provided evidence base for clinical decision making.

In **chapter 7**, we used PCOS as an example, discussing the need of IPD meta-analysis to guide personalised treatment choices and to refine the existing diagnostic criteria.

In **chapter 8**, we conducted an IPD meta-analysis on the first-line ovulation induction strategy for women with PCOS and provided insights into personalised medicine.

In **chapter 9**, provided a summary of the thesis, and discussed directions for future research.

REFERENCES

Broeze KA, Opmeer BC, van der Veen F, Bossuyt PM, Bhattacharya S, Mol BW. Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. *Hum Reprod Update* 2010;16: 561-567.

Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions. In Julian PT Higgins SG (ed). 2011. Cochrane Collaboration.

Inhorn MC, Patrizio P. Infertility around the globe: new thinking on gender, reproductive technologies and global movements in the 21st century. *Hum Reprod Update* 2015;21: 411-426.

Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011;14: 417-428.

Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23: 3105-3124.

Mascarenhas MN, Flaxman SR, Boerma T, Vanderpoel S, Stevens GA. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLoS Med* 2012;9: e1001356.

Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ* 2013;346: f2914.

Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340: c221.

Rockliff HE, Lightman SL, Rhidian E, Buchanan H, Gordon U, Vedhara K. A systematic review of psychosocial factors associated with emotional adjustment in in vitro fertilization patients. *Hum Reprod Update* 2014;20: 594-613.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 2012;3: 80-97.

Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338: b1147.

Thompson SG, Higgins JP. Treating individuals 4: can meta-analysis help target interventions at individuals most likely to benefit? *Lancet* 2005;365: 341-346.

Chapter 2

Network meta-analyses in reproductive medicine: challenges and opportunities

Wang R, Seidler AL, Askie L, Norman RJ, Bhattacharya S, van Wely M, Mol BWJ. Network meta-analyses in reproductive medicine: challenges and opportunities. 2019 Unsubmitted.

Statement of Authorship

Title of Paper	Network meta-analyses in reproductive medicine: challenges and opportunities
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Wang R, Seidler AL, Askie L, Norman RJ, Bhattacharya S, van Wel, M, Mol BWJ. Network meta-analyses in reproductive medicine: challenges and opportunities. 2019 Unpublished

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	RW conceptualised the paper, wrote the first draft and approved the final version.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	2/9/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Anna Lene Seidler		
Contribution to the Paper	ALS conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.		
Signature		Date	04/09/2019

Name of Co-Author	Lisa Askie
Contribution to the Paper	LA conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.

Signature		Date	04/09/2019

Name of Co-Author	Robert J Norman		
Contribution to the Paper	RJN conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.		
Signature		Date	03/09/2019

Name of Co-Author	Siladitya Bhattacharya		
Contribution to the Paper	SB conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.		
Signature		Date	03/09/2019

Name of Co-Author	Madelon van Wely		
Contribution to the Paper	MvW, BWM conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.		
Signature		Date	03/09/2019

Name of Co-Author	Ben Willem Mol		
Contribution to the Paper	BWM conceptualised the paper, critically revised the manuscript for important intellectual content, and approved the final version.		
Signature		Date	2/9/2019

ABSTRACT

Network meta-analyses synthesise both direct and indirect evidence and allow simultaneous comparisons of multiple treatments. Relatively new in reproductive medicine, this approach has gained in popularity and interest in recent years. In this paper, we briefly introduce the principles of network meta-analysis and explain key underlying transitivity assumption. In addition, we present a search for published network meta-analyses in reproductive medicine, summarize their challenges and provide insights into future research opportunities.

KEYWORDS

network meta-analysis, evidence synthesis, reproductive medicine, infertility

INTRODUCTION

Systematic reviews and meta-analyses are an integral part of evidence-based medicine and since they summarize all the available knowledge at a certain point in time, they are essential for guiding both clinical decision-making and planning of future research (Djulbegovic and Guyatt, 2017). Conventional pairwise meta-analyses of randomised controlled trials (RCTs) involve the comparison of two interventions at a time for a specific clinical condition. However, in clinical practice, it is common to find multiple interventions for a specific condition which allow a number of pairwise meta-analyses. For example, ovulation induction in women with WHO II anovulation can be done with metformin, clomiphene citrate (CC), gonadotrophins, letrozole or tamoxifen. In such a clinical scenario, doing multiple pairwise meta-analyses may not directly guide clinical decisions among multiple treatment options. In addition, individual RCTs usually compare an intervention to an existing standard treatment or placebo/ no treatment. Therefore, it is unrealistic to expect all the existing treatment options for a woman with a particular condition such as polycystic ovary syndrome (PCOS) to have been compared with each other in subsequent pairwise meta-analyses due to the unavailability of primary RCTs.

Network meta-analysis, also known as multiple-treatments meta-analysis, allows multiple interventions to be compared and ranked simultaneously based on their summary results (Riley et al., 2017). Its framework combines all evidence from available RCTs and deduces indirect evidence for comparisons that were not studied in a RCT. Subsequently, the relative effectiveness of all interventions can be estimated.

Indirect comparison is a special form of this framework by presenting the indirect evidence of a specific comparison generated from direct evidence via a common comparator (Figure 1a). For instance, take PCOS with three treatment options A, B, and C (Figure 1a, referring to CC, no treatment and letrozole, respectively). There are only studies available comparing treatments A and B, and studies comparing treatments A and C, but no studies comparing B

and C. Indirect comparison would utilise the information from the existing studies, to make an inference about how B compares to C.

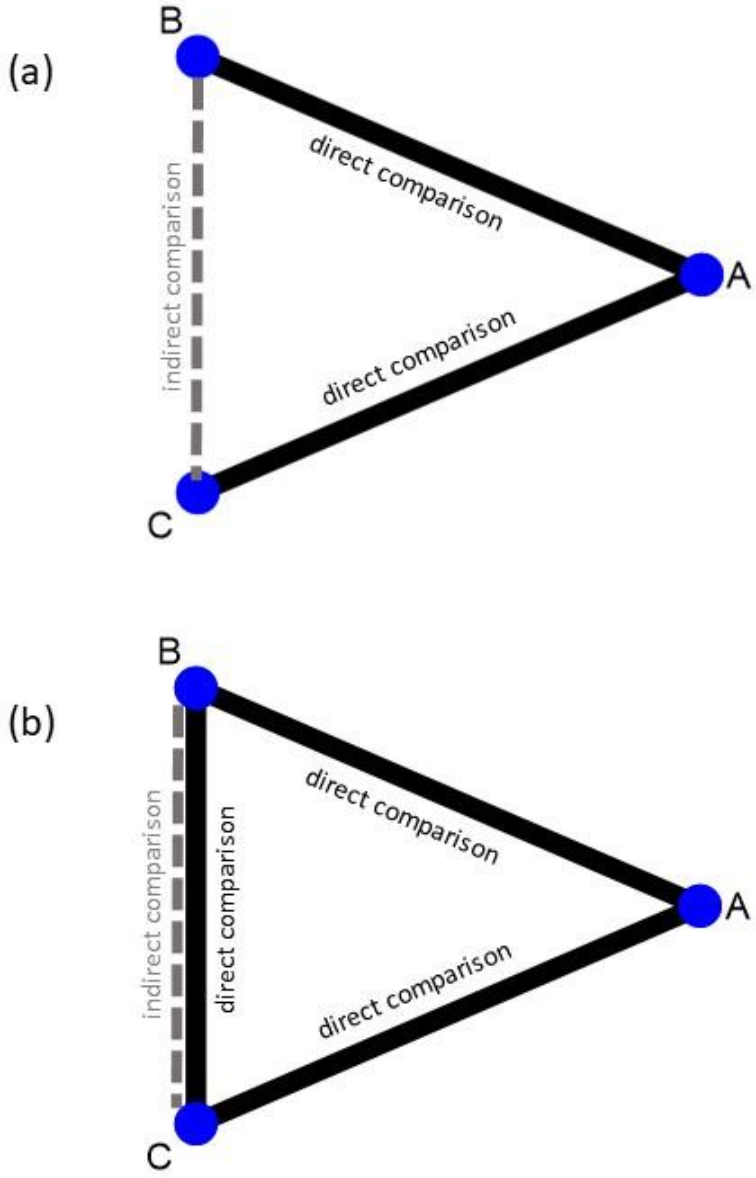


Figure 1 Network plots of three interventions

Network meta-analyses are becoming increasing popular in the literature. A search in PubMed with a combination of the index term "network meta-analysis" and the free word "network meta-analys*" on 6th August 2019 yielded over 3300 results (Figure 2). Figure 2 shows an exponential increase of published network meta-analyses in the medical literature,

implying the increasing applications of network meta-analyses in different clinical areas as well as the methodological developments of network meta-analyses over the last decade.

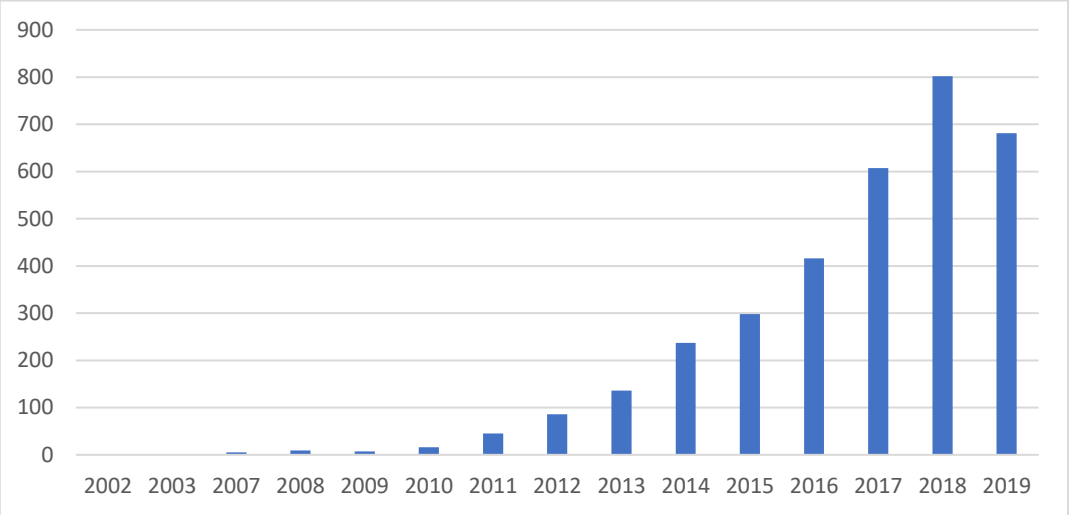


Figure 2 Numbers of publications on network meta-analyses identified in PubMed

KEY ASSUMPTION IN NETWORK META-ANALYSES

The key assumption underlying network meta-analysis is transitivity. The transitivity assumption, also called similarity or exchangeability assumption, means that the comparison between two treatments can be made via a common comparator, i.e. the indirect evidence of B vs C can be estimated by using the direct evidence of B vs A and C vs A (Figure 1b) (Cipriani et al., 2013). Transitivity refers to the clinical and methodological similarities of different comparisons in a network. In Figure 1b, if the transitivity assumption is valid, treatment A should be similar in trials comparing B vs A and C vs A. One conceptual interpretation of the transitivity assumption is that in a theoretical multi-arm RCT, participants could, in principle, be randomized to any of the included interventions, i.e. the interventions are jointly randomisable (Salanti, 2012). Transitivity can also be understood in relation to the distribution of effect modifiers. An effect modification occurs, if a treatment has a different effect among different groups of participants, or in different study settings. In Figure 1b the distribution of effect modifiers should not be substantially different in trials comparing B vs A and C vs A (Jansen and Naci, 2013). For instance, in women with PCOS, if all trials comparing to CC vs letrozole refers to treatment naïve women while all trials

comparing CC vs FSH refers to CC-resistant women, treatment history is an effect modifier and therefore the transitivity assumption is violated here. To make the transitivity assumption valid, we can only include all interventions that have been compared in treatment-naïve women in a network meta-analysis.

It is worth noting that similarities should be only applicable to effect modifiers, not prognostic factors, as prognostic factors that affect all treatment-arms in a similar way do not alter relative treatment effects, i.e., treatment effects in populations with different prognostic factors are the same In contrast, effect modifiers have impact on the prognosis in one arm, but not or to a lesser extend in another arm. As transitivity is a conceptual definition, it may be difficult to evaluate statistically. A common approach is to compare the distribution of potential effect modifiers across different sets of trials in different comparisons within a network (Jansen and Naci, 2013; Salanti, 2012). The transitivity assumption should be considered when defining the research question, as generally broad research questions are more likely to violate the transitivity assumptions. For instance, if the research question is to compare all treatment strategies for polycystic ovary syndrome (PCOS) including IVF, the transitivity assumption will be violated. IVF is the third-line choice for women with PCOS after ovulation induction with CC, letrozole, gonadotrophins or even ovarian drilling, and therefore the participants in trials comparing IVF vs other treatments should have failed ovulation induction. Joint randomisation is impossible in this case, as treatment-naïve women could never be randomized to the IVF group. In light of the transitivity assumption, narrowing the research questions into first-line ovulation induction treatments only can solve the problem and is more useful in clinical decision-making (Wang et al., 2017).

Consistency, also known as coherence, refers to the agreement between direct and indirect evidence for the same comparison. Consistency is a statistical manifestation of transitivity and can only be evaluated when both direct and indirect evidence are available in the same comparison, i.e. in a closed loop (Cipriani et al., 2013; Salanti, 2012). Figure 1b shows an

example of a closed loop, in which all treatments have been compared to each other directly. If in the BC comparison in Figure 1b, the direct and indirect evidence yield substantially a different results, inconsistency is observed. For instance, in comparisons B vs A and A vs C, B is better than A and A is better than C, therefore the indirect evidence shows B is better than C. However, if C is better than B in the direct comparison between C vs B, we call it inconsistency.

There are different statistical methods to assess such an inconsistency, including global and local approaches. In a global approach, consistency is tested in a network as a whole and a design-by-treatment interactions model is commonly used (Higgins et al., 2012). In a local approach, consistency is tested in a specific closed loop (loop-specific approach) or in a specific comparison (side-splitting or node-splitting method) by comparing the direct and indirect evidence (Dias et al., 2010; Higgins et al., 2012). Absence of inconsistency does not necessarily mean that the transitivity assumption is valid. However, if inconsistency is detected, the transitivity assumption is no longer valid.

In a pairwise meta-analysis, clinical and methodological heterogeneity should be evaluated before testing statistical heterogeneity. Similar principles should be applied to a network meta-analysis: transitivity assumption should be evaluated before testing consistency. Strategies to deal with detected inconsistency include exploring potential sources of inconsistency by subgroup analyses or meta-regression, using inconsistency models or avoiding performing network meta-analyses (Cipriani et al., 2013).

REPORTING STANDARD

Similar to other meta-analyses, network meta-analysis should be based on a pre-specified protocol. A protocol for a network meta-analysis has additional requirements (in the context of the above-mentioned key assumptions) which are crucial for developing the research question, planning the search and conducting the analysis. (Chaimani et al., 2017)

The PRISMA Extension Statement for Reporting of Systematic Reviews Incorporating Network Meta-analyses of Health Care Interventions has been developed to guide the reporting of network meta-analysis (Hutton et al., 2015). In addition to standard PRISMA guidelines for pairwise meta-analyses, additional information on network geometry, inconsistency assessment and ranking should be reported.

Two frameworks are available to summarise the overall certainty of evidence in network meta-analyses, both of which endorse the principle of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) by incorporating risk of bias, imprecision, inconsistency, indirectness and publication bias into the assessment of each comparison (Puhan et al., 2014; Salanti et al., 2014). In the Puhan approach, when both direct and indirect evidence are available for a particular comparison, the higher of the two certainty ratings is used for the overall rating (Puhan et al., 2014), while in the Salanti approach, the overall certainty rating is based on the contribution of each study to the results from network meta-analysis, allowing a more accurate evaluation (Salanti et al., 2014). The latter has been implemented in a web-based tool - the Confidence In Network Meta-Analysis (CINeMA) tool (Nikolakopoulou et al., 2019; Salanti et al., 2014).

CURRENT STATUS OF NETWORK META-ANALYSES IN REPRODUCTIVE MEDICINE

We searched PubMed from inception up to 6th August 2019 for published network meta-analyses or protocols in reproductive medicine, with network meta-analysis and infertility/fertility/pregnancy/live birth as key search terms (Supplemental Table 1). We also searched the Cochrane Gynaecology and Fertility Group and the Cochrane Pregnancy and Childbirth Group for published network meta-analyses or protocols on topics in reproductive medicine and identified 96 hits. After excluding studies that were not network meta-analyses or studies in other areas, we found 15 network meta-analyses in reproductive medicine

(Abou-Setta, 2006; Abou-Setta, 2007; Al Wattar et al., 2019; Chen et al., 2019; Guo et al., 2016; Lv et al., 2018; Simopoulou et al., 2019; Song et al., 2019; Tsiami et al., 2016; Wang et al., 2017; Wang et al., 2019; Wu et al., 2017; Yan and Xu, 2018; Yu et al., 2017; Zhang et al., 2015) and four Cochrane protocols for network meta-analyses in this research area (Dong et al., 2019; Gallos et al., 2017a; Gallos et al., 2017b; Tjon-Kon-Fat et al., 2017).

Characteristics of these network meta-analyses and protocols are presented in Table 1. These papers were published between 2006 and 2019, with the majority (13 network meta-analyses and four protocols) published after 2015. All papers were in English except one in Chinese (Song et al., 2019). Over 60% of the network meta-analyses or protocols (11/18) were published in general medical journals. The research questions involve comparing treatment options for PCOS, miscarriage, recurrent miscarriage, endometriosis, unexplained infertility, hydrosalpinx and uterine adhesion, controlled ovarian stimulation protocols and embryo transfer techniques during IVF, as well as the choices of contrast media during tubal testing. The number of included studies varied between 2 and 57 studies, the number of participants between 314 and 9250, while the number of interventions compared in one network meta-analysis varied between 3 and 15. Only one third (5/15) of the published network meta-analyses were based on a pre-specified protocol, of which one protocol did not specify the plan to conduct a network meta-analysis (Wu et al., 2017) and one protocol was not accessible (Guo et al., 2016).

The majority (10/18) used frequentist approaches to conduct network meta-analyses. Two network meta-analyses also included cohort studies, but no additional statistical considerations had been used to address the inclusion of non-randomised studies (Chen et al., 2019; Simopoulou et al., 2019). With regards to the reporting standard, only 30% (4/13) followed the PRISMA extension statement for network meta-analysis (Al Wattar et al., 2019; Tsiami et al., 2016; Wang et al., 2017; Wang et al., 2019).

Prospective registration of systematic reviews promotes transparency, helps reduce potential for bias and serves to avoid unintended duplication of reviews (Stewart et al., 2012). Systematic reviews with registered protocols have been shown to be associated with increased quality (Sideri et al., 2018). A pre-specified protocol is even more important in a network meta-analysis, as the key additional assumptions need to be considered so that a relevant search and analysis plan can be developed (Chaimani et al., 2017). Unfortunately, protocol registration has been infrequent overall in the published network meta-analyses in reproductive medicine. Consequently, key assumption on transitivity for network meta-analyses were not considered carefully in some publications.

For instance, one network meta-analysis intended to include all relevant treatments for endometriosis, but the search strategy only covered a few interventions (Chen et al., 2019). In another example, all western medicine was considered as a single intervention in a network meta-analysis on PCOS, which violated the transitivity assumption (Song et al., 2019).

The number of published network meta-analyses has increased over the past several years and is expected to increase further in the future. The main challenges are similar to those in pairwise meta-analyses (ESHRE Capri Workshop Group, 2018; van Wely, 2014), including conduct, reporting and quality of primary studies. The majority of published network meta-analyses in reproductive medicine did not have a pre-specified protocol or follow the reporting guideline in the final report. In addition, the key assumption on transitivity have not been addressed entirely satisfactorily so far. These will challenge the credibility of the results and implementation in clinical practice.

Table 1 Characteristics of network meta-analyses and protocols in reproductive medicine

Publication	Journal	Participants	Type of included studies	Interventions	Number of studies included	Number of interventions and comparators	Protocol	Number of participants	PRISMA-NMA statement cited
Abou-Setta 2006	Reprod Biomed Online	Women undergoing embryo transfer	RCTs	Embryo transfer with different catheters	2	3	no	314	/
Abou-Setta 2007	Reprod Biomed Online	Women undergoing embryo transfer	RCTs	Embryo transfer with different fundus-to-catheter distances	3	3	no	2170	/
Zhang 2015	Medicine	Women with recurrent miscarriage	RCTs	Antithrombotic treatments	19 (12)	4	no	2391	No
Guo 2016	Sci Rep	Women undergoing controlled ovarian stimulation	RCTs	Pharmacologic therapies for the prevention of OHSS	31	11	Not accessible	7181	No
Tsiami 2016	Ultrasound Obstet Gynecol	Women with hydrosalpinx	RCTs	Hydrosalpinx treatments prior to IVF-ET	7	4	no	859	Yes
Wang 2017	BMJ	Women with WHO II anovulation	RCTs	Ovulation induction strategies	57	8	yes	8082	Yes
Wu 2017	Sci Rep	Women with missed miscarriage	RCTs	Different routes of administration of misoprostol	18	9	Yes, not planned as a NMA	1802	No
Yu 2017	Sci Rep	Women with clomiphene citrate – resistant PCOS	RCTs	Ovulation induction strategies	26	9	No	2722	No

Lv 2018	Am J Reprod Immunol	Women with recurrent miscarriage	RCTs	Various hormone therapies, immunotherapies, and antithrombotic therapies	49	15	No	8496	No
Yan 2018	J Minim Invasive Gynecol	Women with intrauterine adhesions or with high risk of intrauterine adhesions	RCTs	Adjuvant interventions including medical interventions, intrauterine devices and barriers gels	20	10	No	1891	No
Al Wattar 2019	Hum Reprod Update	Women with first-trimester miscarriage	RCTs	Expectant, medical, placebo, surgical and a combination of any medical plus surgical treatment	46	7	Yes	9250	Yes
Chen 2019	J Cell Biochem	Women with endometriosis	RCTs and cohort studies	Different moxibustion, Chinese herbal medicine, moxibustion approached and western medicine or placebo	10	8	No	1263	No
Simopoulou 2019	J Assist Reprod Genet	Women undergoing IVF/ICSI	RCTs and cohort studies	day-2 (D2), day-3 (D3), and day-5 (D5) ET	15	4	No	3319	No
Song 2019	Chinese acupuncture & moxibustion	Women with PCOS	RCTs		39	14	No	4605	No
Wang 2019	Ultrasound Obstet Gynecol	Women with infertility undergoing tubal patency test	RCTs	Tubal flushing with different contrast media	14	4	yes	3852	Yes

Gallos 2017 (protocol)	Cochrane Database of Systematic Reviews	Women undergoing controlled ovarian stimulation protocols	RCTs	Different ovarian stimulation protocols	/	/	/	/	/
Gallos 2017 (protocol)	Cochrane Database of Systematic Reviews	Women with miscarriage	RCTs	Managements of miscarriage	/	/	/	/	/
Tjon-Kon-Fat 2017 (protocol)	Cochrane Database of Systematic Reviews	Couples with unexplained infertility	RCTs	Interventions including expectant management, ovarian stimulation, IUI, IUI with ovarian stimulation, IVF/ICSI	/	/	/	/	/
Dong 2019 (protocol)	Medicine	Women with endometriosis	RCTs	Different traditional Chinese patent medicine	/	/	/	/	

OPPORTUNITIES

Network meta-analyses are expected to be extremely helpful for clinical conditions where decisions need to be made among multiple treatment options, especially in the areas where novel techniques are emerging. In reproductive medicine, there are many clinical conditions where decision-making needs to be based on multiple treatment strategies, including for example, the clinical managements of PCOS, unexplained infertility and idiopathic oligo-astheno-teratozoospermia, but also ectopic pregnancy. In addition, IVF is a multi-step process including controlled ovarian stimulation, oocytes collection, insemination, embryo culture, selection, freezing and transfer and luteal phase support, where each step involves a decision among multiple choices including novel techniques. The application of network-meta-analyses would and should be helpful to answer these research questions.

CONCLUSIONS

Network meta-analyses provide opportunities for simultaneous comparison of multiple interventions and their use represents a promising strategy in guiding clinical decision-making in reproductive medicine. However, the validity of their findings depends on clarity around assumptions regarding transitivity and consistency. Additional assumptions on transitivity and consistency are important to guarantee the validity of the findings in network meta-analyses for clinical practice and the scheduling of research priorities. However, the majority of existing network meta-analyses in reproductive medicine suffer from suboptimal conduct and report. Improvements in the methodology are expected to boost our confidence in the clinical implications of findings derived from network-meta-analysis.

REFERENCES

Abou-Setta AM. Firm embryo transfer catheters for assisted reproduction: a systematic review and meta-analysis using direct and adjusted indirect comparisons. *Reprod Biomed Online* 2006;12: 191-198.

Abou-Setta AM. What is the best site for embryo deposition? A systematic review and meta-analysis using direct and adjusted indirect comparisons. *Reprod Biomed Online* 2007;14: 611-619.

Al Wattar B, Murugesu N, Tobias A, Zamora J, Khan KS. Management of first-trimester miscarriage: a systematic review and network meta-analysis. *Hum Reprod Update* 2019;25: 362-374.

Chaimani A, Caldwell DM, Li T, Higgins JPT, Salanti G. Additional considerations are required when preparing a protocol for a systematic review with multiple interventions. *J Clin Epidemiol* 2017;83: 65-74.

Chen Y, Wang H, Wang S, Shi X, Wang Q, Ren Q. Efficacy of ten interventions for endometriosis: A network meta-analysis. *J Cell Biochem* 2019;120: 13076-13084.

Cipriani A, Higgins JP, Geddes JR, Salanti G. Conceptual and technical challenges in network meta-analysis. *Ann Intern Med* 2013;159: 130-137.

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

Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet* 2017;390: 415-423.

Dong S, Zhang J, Zhai F, Zhao X, Qin X. Comparative efficacy and safety of traditional Chinese patent medicine for endometriosis: A Bayesian network meta-analysis protocol. *Medicine (Baltimore)* 2019;98: e16473.

ESHRE Capri Workshop Group. Protect us from poor-quality medical research. *Hum Reprod* 2018;33: 770-776.

Gallos ID, Eapen A, Price MJ, Sunkara SK, Macklon NS, Bhattacharya S, Khalaf Y, Tobias A, Deeks JJ, Rajkhowa M *et al.* Controlled ovarian stimulation protocols for assisted reproduction: a network meta‐analysis. *Cochrane Database of Systematic Reviews* 2017a.

Gallos ID, Williams HM, Price MJ, Eapen A, Eyo MM, Tobias A, Deeks JJ, Tunçalp Ö, Gülmezoglu AM, Coomarasamy A. Methods for managing miscarriage: a network meta‐analysis. *Cochrane Database of Systematic Reviews* 2017b.

Guo JL, Zhang DD, Zhao Y, Zhang D, Zhang XM, Zhou CQ, Yao SZ. Pharmacologic Interventions in Preventing Ovarian Hyperstimulation Syndrome: A Systematic Review and Network Meta-Analysis. *Sci Rep* 2016;6: 19093.

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. *Res Synth Methods* 2012;3: 98-110.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162: 777-784.

Jansen JP, Naci H. Is network meta-analysis as valid as standard pairwise meta-analysis? It all depends on the distribution of effect modifiers. *BMC Med* 2013;11: 159.

Lv S, Yu J, Xu X. A comparison of effectiveness among frequent treatments of recurrent spontaneous abortion: A Bayesian network meta-analysis. *Am J Reprod Immunol* 2018;80: e12856.

Nikolakopoulou A, Higgins JP, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, Salanti G. Assessing Confidence in the Results of Network Meta-Analysis (CINeMA). *bioRxiv* 2019: 597047.

Puhan MA, Schunemann HJ, Murad MH, Li T, Brignardello-Petersen R, Singh JA, Kessels AG, Guyatt GH, Group GW. A GRADE Working Group approach for rating the quality of treatment effect estimates from network meta-analysis. *Bmj* 2014;349: g5630.

Riley RD, Jackson D, Salanti G, Burke DL, Price M, Kirkham J, White IR. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *Bmj* 2017;358: j3932.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 2012;3: 80-97.

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.

Sideri S, Papageorgiou SN, Eliades T. Registration in the international prospective register of systematic reviews (PROSPERO) of systematic review protocols was associated with increased review quality. *J Clin Epidemiol* 2018;100: 103-110.

Simopoulou M, Sfakianoudis K, Tsioulou P, Rapani A, Maziotis E, Giannelou P, Grigoriadis S, Pantou A, Nikolettos K, Vlahos N *et al.* Should the flexibility enabled by performing a day-4 embryo transfer remain as a valid option in the IVF laboratory? A systematic review and network meta-analysis. *J Assist Reprod Genet* 2019;36: 1049-1061.

Song YJ, Liang FX, Wu S, Yang HS, Chen L, Huang Q, Tang HT, Lu W, Wang H, Chen S *et al.* [Network meta-analysis on the effects of the acupuncture-related therapy on ovulation rate and pregnancy rate in patients with polycystic ovary syndrome]. *Zhongguo Zhen Jiu* 2019;39: 792-798.

Stewart L, Moher D, Shekelle P. Why prospective registration of systematic reviews makes sense. *Syst Rev* 2012;1: 7.

Tjon-Kon-Fat RI, Wang R, Eijkemans MJC, Bossuyt PMM, Mol BWJ, Van der Veen F, Bhattacharya S, van Wely M. Interventions for unexplained subfertility: a systematic review and network meta‐analysis. *Cochrane Database of Systematic Reviews* 2017.

Tsiami A, Chaimani A, Mavridis D, Siskou M, Assimakopoulos E, Sotiriadis A. Surgical treatment for hydrosalpinx prior to in-vitro fertilization embryo transfer: a network meta-analysis. *Ultrasound Obstet Gynecol* 2016;48: 434-445.

van Wely M. The good, the bad and the ugly: meta-analyses. *Hum Reprod* 2014;29: 1622-1626.

Wang R, Kim BV, van Wely M, Johnson NP, Costello MF, Zhang H, Ng EH, Legro RS, Bhattacharya S, Norman RJ *et al.* Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis. *Bmj* 2017;356: j138.

Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, Mijatovic V, Mol BW. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. *Ultrasound Obstet Gynecol* 2019;54: 172-181.

Wu HL, Marwah S, Wang P, Wang QM, Chen XW. Misoprostol for medical treatment of missed abortion: a systematic review and network meta-analysis. *Sci Rep* 2017;7: 1664.

Yan Y, Xu D. The Effect of Adjuvant Treatment to Prevent and Treat Intrauterine Adhesions: A Network Meta-Analysis of Randomized Controlled Trials. *J Minim Invasive Gynecol* 2018;25: 589-599.

Yu Y, Fang L, Zhang R, He J, Xiong Y, Guo X, Du Q, Huang Y, Sun Y. Comparative effectiveness of 9 ovulation-induction therapies in patients with clomiphene citrate-resistant polycystic ovary syndrome: a network meta-analysis. *Sci Rep* 2017;7: 3812.

Zhang T, Ye X, Zhu T, Xiao X, Liu Y, Wei X, Liu Y, Wu C, Guan R, Li X *et al.* Antithrombotic Treatment for Recurrent Miscarriage: Bayesian Network Meta-Analysis and Systematic Review. *Medicine (Baltimore)* 2015;94: e1732.

Chapter 3

Treatment strategies for women with WHO group II anovulation - systematic review and network meta-analysis

Statement of Authorship

Title of Paper	Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Wang R, Kim BV, van Wely M, Johnson NP, Costello MF, Zhang H, Ng EH, Legro RS, Bhattacharya S, Norman RJ, Mol BW. Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis. BMJ. 2017 Jan 31;356:j138. doi: 10.1136/bmj.j138.

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	RW contributed to the study conception and design, interpreted the work, drafted the manuscript, commented on the drafts, approved the final draft and acted as corresponding author.		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	21/2/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

iv. the candidate's stated contribution to the publication is accurate (as detailed above);

v. permission is granted for the candidate in include the publication in the thesis; and

vi. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Bobae V Kim		
Contribution to the Paper	BVK collected the data, interpreted the work, drafted the manuscript, commented on the drafts and approved the final draft.		
Signature		Date	1/3/17

Name of Co-Author	Madelon van Wely
Contribution to the Paper	MVW contributed to the study conception, interpreted the work, provided technical support, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.

Signature		Date	22/2/2017

Name of Co-Author	Neil P Johnson		
Contribution to the Paper	NPJ contributed to the study conception, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	23/2/2017

Name of Co-Author	Michael F Costello		
Contribution to the Paper	MFC contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	22/2/2017

Name of Co-Author	Hanwang Zhang		
Contribution to the Paper	HZ contributed to the study conception, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	24/2/2017

Name of Co-Author	Ernest Hung Yu Ng		
Contribution to the Paper	EHYN interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	22/2/2017

Name of Co-Author	Richard S Legro		
Contribution to the Paper	RSL interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	27/2/2017

Name of Co-Author	Siladitya Bhattacharya
Contribution to the Paper	SB interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.

Signature		Date	22/2/2017

Name of Co-Author	Robert J Norman		
Contribution to the Paper	RJN contributed to the study conception, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	22/2/2017

Name of Co-Author	Ben Willem J Mol		
Contribution to the Paper	BWJM contributed to the study conception, collected the data, analysed the data, interpreted the work, drafted the manuscript, commented on the drafts and approved the final draft.		
Signature		Date	22/2/2017

ABSTRACT

Objective: To compare the effectiveness of alternative first-line treatment options in women with WHO group II anovulation wishing to conceive.

Design: Systematic review and network meta-analysis.

Data sources: Cochrane Central Register of Controlled Trials, MEDLINE and EMBASE.

Study selection: Randomised controlled trials comparing eight ovulation induction treatments in women with WHO group II anovulation: clomiphene, letrozole, metformin, combined clomiphene-metformin, tamoxifen, gonadotropins, laparoscopic ovarian drilling and placebo/no treatment. We assigned study quality utilizing the methodology and categories described in the Cochrane Collaboration Handbook. We chose pregnancy, defined preferably as clinical pregnancy, as the primary outcome. Live birth, ovulation, miscarriage and multiple pregnancy were secondary outcomes.

Results: Of the 2,631 titles and abstracts initially identified, we included 57 randomised controlled trials reporting on 8,082 women with WHO group II anovulation. All pharmacological treatments were superior to placebo or no intervention in terms of pregnancy and ovulation. Compared to clomiphene, both letrozole and the combination of clomiphene and metformin showed higher pregnancy rates (odds ratio 1.53, 95% confidence interval 1.25 to 2.85 and odds ratio 1.56, 95% confidence interval 1.24 to 1.97) and ovulation rates (odds ratio 1.99, 95% confidence interval 1.38 to 2.87 and odds ratio 1.55, 95% confidence interval 1.02 to 2.36, respectively). Letrozole led to higher live birth rates than clomiphene alone (odds ratio 1.67, 95% confidence interval 1.11 to 2.49). Both letrozole (odds ratio 0.46, 95% confidence interval 0.23 to 0.92) and metformin (odds ratio 0.22, 95% confidence interval 0.05 to 0.92) led to lower multiple pregnancy rates than clomiphene alone.

Conclusions: In women with WHO group II anovulation, letrozole and the combination of clomiphene and metformin are superior to clomiphene alone in terms of ovulation and

pregnancy. Letrozole is the only therapy showing a statistically significantly higher live birth rate than clomiphene alone.

Systematic review registration: PROSPERO CRD42015027579

What is already known on this topic?

- Clomiphene is the long standing first-line treatment for WHO group II anovulation.
- Existing pairwise meta-analyses are limited to comparisons of two treatments.

What this study adds?

- This is the first study to compare all the most common ovulation induction regimens with each other, using direct and indirect means.
- All pharmacological ovulation inductions are superior to placebo/no treatment in terms of ovulation and pregnancy in women with WHO group II anovulation,
- Letrozole is the most effective treatment in terms of live birth, and one of the top 3 treatments in terms of pregnancy and ovulation.
- The combination of clomiphene and metformin is the most effective treatment in terms of pregnancy, but not live birth. The potential higher chances of side effects should also be taken into account in decision making.
- Metformin and letrozole are associated with the lowest rates of multiple pregnancy.

INTRODUCTION

Infertility affects one in seven couples and ovulation disorders account for a quarter of all cases (National Institute for Health and Care Excellence, 2013). Normogonadotrophic anovulation, also classified as World Health Organization (WHO) group II anovulation, is the most common category of anovulatory infertility and within this group polycystic ovary syndrome (PCOS) is by far the most prevalent cause (ESHRE Capri Workshop Group, 2012). PCOS was first described in 1935 by Stein and Leventhal (Stein and Leventhal, 1935). Previously described in a number of different ways, the diagnostic criteria for PCOS, agreed jointly by the European Society of Human Reproduction and Embryology and the American Society for Reproductive Medicine, are known as the Rotterdam criteria (Rotterdam ESHRE ASRM-Sponsored PCOS Consensus Workshop Group, 2004a, b). These criteria are also endorsed by the Endocrine Society (Legro et al., 2013) and are used by a wide range of medical professionals, not just obstetricians and gynaecologists. The clinical manifestations of PCOS include oligomenorrhea or amenorrhea, hirsutism, and frequently infertility (Sirmans and Pate, 2013). When women with PCOS conceive, they and their infants are at increased risk of perinatal complications, including gestational diabetes, pre-eclampsia, preterm labor and neonatal morbidity (Amsterdam ESHRE ASRM-Sponsored 3rd PCOS Consensus Workshop Group, 2012; Boomsma et al., 2006; Fauser et al., 2012).

Safe and effective ovulation induction is important for women with WHO group II anovulation who wish to conceive, to avoid premature exposure to in-vitro fertilisation, which is invasive, expensive and associated with potentially higher chances of perinatal complications and congenital abnormalities (Hansen et al., 2013; Hart and Norman, 2013; Pandey et al., 2012; Pinborg et al., 2013). A number of medical options are used to treat women with ovulation disorders suffering from infertility, including oestrogen receptor modulators (such as clomiphene and tamoxifen), aromatase inhibitors (such as letrozole), insulin-sensitizing drugs (such as metformin), and direct hormonal stimulation of the ovaries (gonadotropins), with laparoscopic ovarian drilling being a surgical alternative.

Traditional pairwise meta-analysis only allows comparison of two ovulation induction interventions (Brown et al., 2009; Franik et al., 2014; Misso et al., 2013; Misso et al., 2012; Moll et al., 2007; Tang et al., 2012). However, many of these treatment strategies have not been compared directly in previous randomised controlled trials. Therefore, it is difficult to identify the most effective treatment based on direct evidence. Network meta-analysis, also known as multiple treatment comparison meta-analysis, allows the comparisons of multiple treatments in a single statistical model (Lu and Ades, 2004; Mills et al., 2013; Song et al., 2009), and a hierarchy of effectiveness of these treatments that can guide decision making (Jansen et al., 2011; Salanti, 2012). The application of network meta-analysis is crucial in areas where multiple interventions are available, such as in WHO group II anovulation.

We therefore performed a systematic review and network meta-analysis to compare the effectiveness of different treatment options, including clomiphene, letrozole, metformin, combined clomiphene-metformin, tamoxifen, gonadotropins, laparoscopic ovarian drilling and placebo/no treatment, in women with WHO group II anovulation and to identify the best first-line treatment strategy (Systematic review registration: PROSPERO CRD42015027579).

METHODS

Search strategy and selection criteria

We conducted and reported the study according to the PRISMA extension statement for network meta-analyses (Hutton et al., 2015). We performed an extensive electronic search of the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, and EMBASE for randomised controlled trials. The search strategies were based on combinations of ovulation induction and anovulation (or PCOS), using both free words and index terms (Appendix 1). We sought further trial details or protocols to establish eligibility of potential trials. We also searched previous published Cochrane systematic reviews on ovulation induction for additional studies. No language restrictions were applied. Our latest search was completed on April 11th, 2016.

We included published and unpublished randomised controlled trials comparing one or more common ovulation induction options with placebo, no treatment or other treatments: clomiphene, tamoxifen, letrozole, metformin, gonadotropins, laparoscopic ovarian drilling or the combination of clomiphene and metformin. Treatments were categorized according to the initial randomised allocation, although subsequent clinical management may have included further doses or an alternative treatment.

Studies were excluded if they were not randomised controlled trials, only included treatment resistant women or failed to report on clinical pregnancy, live birth or pregnancy. The population within the included studies was classified as: (1) treatment naïve women, (2) a combination of treatment naïve and treatment exposed women, and (3) women whose treatment status was unknown. Crossover trials were also included if pre-cross over data were available. Studies were also excluded if they only compared different doses of the same treatment option or compared the effects of adding medical adjuncts such as dexamethasone. Authors were contacted for further information if necessary.

Patient involvement

There was no patient involvement in framing the research question, choosing the outcome measures or conducing the research. We plan to involve Fertility Network UK, PCOS Challenge, RESOLVE and Access Australia's National Infertility Network Ltd in the dissemination of the research results by means of short, easy to read summaries of key results, infographics and audio or video interviews that can be used by patients and caregivers.

Data extraction and assessment of risk of bias

Two reviewers (R.W. and B.V.K) independently assessed the eligibility of all identified citations, and extracted data from original trial reports using a specifically designed form capturing information on study design, trial setting, patient characteristics (inclusion criteria, age, body mass index, duration of infertility, history of ovulation induction), sample sizes, details of ovulation induction options, and outcomes. Disagreements were referred to a third reviewer (B.W.J.M) to reach consensus.

We chose pregnancy, defined preferably as clinical pregnancy, as the primary outcome. Clinical pregnancy was defined as either pregnancy visualized at ultrasonography of one or more gestational sacs (Harbin Consensus Conference Workshop Group, 2014; Harbin Consensus Conference Workshop Group et al., 2014). Since the comparison of the effectiveness of a treatment based on either clinical pregnancy or live birth rate as endpoints results often in comparable conclusions (Clarke et al., 2010), we used data on live birth or pregnancy (positive human chorionic gonadotropin blood or urine test) as outcome when data on clinical pregnancy were not available. Secondary outcomes were live birth, ovulation, miscarriage and multiple pregnancy.

Study quality was assigned by two reviewers (R.W. and B.V.K) utilizing the methodology and categories described in the Cochrane Collaboration Handbook (Higgins, 2011). Again, in case of disagreement a third reviewer (B.W.J.M) was asked to reach consensus. Briefly, the tool for assessing risk of bias addresses seven specific domains: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of

outcome assessment, incomplete outcome data, selective reporting, and other sources of bias. Each domain is assigned a judgment relating to the risk of bias for that study classified as low risk, high risk or unclear. We presented risk of bias graph by Review Manager 5.3 software (Higgins, 2011).

Data synthesis and statistical analysis

A network meta-analysis was conducted to simultaneously compare seven ovulation induction treatment options and placebo or no treatment for each outcome. In its simplest form, a network meta-analysis is the combination of direct and indirect estimates of relative treatment effect in a single analysis. An indirect estimate of the relative treatment effect A versus B can be formed by comparing direct trials of A versus C with trials of B versus C. Network plots were constructed to illustrate the geometry of the network (Chaimani et al., 2013).

All network meta-analyses were conducted within a random effects multiple regression model using "mvmeta" package in Stata software (Version 12.0, Stata Corp, College Station, TX) (Chaimani et al., 2013; Chaimani and Salanti, 2015). Where direct data were available, pairwise meta-analyses in random effects model were also performed in Stata and the agreement of direct and indirect evidence was assessed by constructing an inconsistency plot. Studies with 0 or 100% events in all interventions were excluded from the analysis because these studies do not allow conclusions on relative effects. For studies with a 0 event in one arm only, a continuity correction of 0.5 was added to each cell. To avoid double counting of events, multi-intervention trials were analyzed in their original form without the need to combine interventions.

We presented network meta-analysis summary treatment effects (odds ratios) with their 95% confidence intervals as well as predictive intervals to facilitate interpretation of the results in the light of the magnitude of heterogeneity (Chaimani et al., 2013). Predictive intervals can provide an interval within which the estimate of a future study is expected to be

(Chaimani et al., 2013). We applied the comparison adjusted funnel plot to assess small study effects in the network. We used the surface under the cumulative ranking curve to rank the treatments (Chaimani et al., 2013; Salanti et al., 2011). It is a percentage of the effectiveness of every treatment relative to an imaginary treatment that is always the best without uncertainty. We then performed sensitivity analysis to explore important network inconsistency. We restricted the analysis to those trials on treatment naïve women, trials with low risk of randomization and allocation bias, and trials reporting clinical pregnancy for sensitivity analysis.

RESULTS

Characteristics of included studies

The literature search yielded a total of 2,631 publications, as is shown in the PRISMA flowchart (Figure 1). Fifty-six(Abuelghar et al., 2013; Amer et al., 2015; Amer et al., 2009; Atay et al., 2006; Ayaz et al., 2013; Aygen et al., 2007; Badawy et al., 2009; Badawy and Gibreal, 2011; Basirat et al., 2012; Bayar et al., 2006; Beigi, 2006; Boostanfar et al., 2001; Boudhraa et al., 2010; Cudmore and Tupper, 1966; Dasari and Pranahita, 2009; Dehbashi et al., 2009; El-Biely and Habba, 2001; Fleming et al., 2002; Garcia et al., 1985; Homburg et al., 2012; Jahan, 2015; Johnson et al., 1966; Johnson et al., 2010; Kar, 2012; Kar and Sanchita, 2015; Karimzadeh et al., 2007; Karimzadeh and Javedani, 2010; Keikha and Shahraki Mojahed, 2011; Khorram et al., 2006; Leanza et al., 2014; Legro et al., 2007; Legro et al., 2014; Liu et al., 2015; Lopez et al., 2004; Lord et al., 2006; Lorzadeh et al., 2011; Maged et al., 2015; Mobusher, 2014; Moll et al., 2006; Nazik and Kumtepe, 2012; Palomba et al., 2005; Raja et al., 2005; Ray et al., 2012; Robinson et al., 2003; Roy et al., 2012; Sahin et al., 2004; Santonocito et al., 2009; Selim and Borg, 2012; Seyedoshohadaei et al., 2012; Sharief and Nafee, 2015; Sheikh-El-Arab and Elmaghraby, 2011; Tang et al., 2006; Vegetti et al., 1999; Williams et al., 2009; Zain et al., 2009; Zeinalzadeh et al., 2010) publications reporting on 57 trials fulfilled the eligibility criteria, as one study (Johnson et al., 2010) included two individual trials (Appendix 2). Five studies (Amer et al., 2015; Amer et al., 2009; Cudmore and Tupper, 1966; Garcia et al., 1985; Lopez et al., 2004) were crossover studies and eight studies (Amer et al., 2015; Beigi, 2006; Jahan, 2015; Keikha and Shahraki Mojahed, 2011; Liu et al., 2015; Robinson et al., 2003; Vegetti et al., 1999; Williams et al., 2009) were reported in conference abstracts. Publication dates ranged from 1966 to 2015, with 45 trials published in the last 10 years. The studies were conducted in a variety of countries. Four studies were reported in French (Boudhraa et al., 2010), Italian (Santonocito et al., 2009), Turkish (Aygen et al., 2007) and Persian (Lorzadeh et al., 2011), respectively.

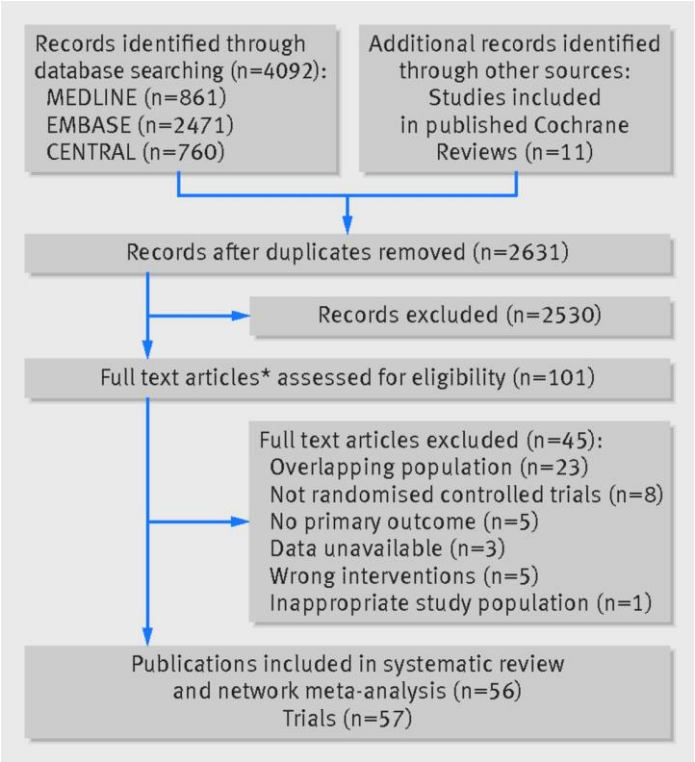


Fig 1 PRISMA flow diagram of literature search for randomised controlled trials comparing eight ovulation induction treatments in women with WHO group II anovulation.
*Full text articles=including abstract only publications

Out of the 57 trials, seven (Jahan, 2015; Johnson et al., 2010; Kar and Sanchita, 2015; Karimzadeh and Javedani, 2010; Legro et al., 2007; Seyedoshohadaei et al., 2012; Zain et al., 2009) had three comparison interventions while each of the remaining 50 trials had two. Overall, 8,082 women with WHO group II anovulation were randomised to seven different treatment options including clomiphene, letrozole, metformin, combined clomiphene-metformin, tamoxifen, gonadotropins and laparoscopic ovarian drilling, and to placebo/no treatment. The network plots are presented in appendix 5 for pregnancy, live birth, ovulation, miscarriage and multiple pregnancy.

Risk of bias assessment results

There were 31 (54%) randomised controlled trials with low risk of bias on random sequence generation and 25 (44%) randomised controlled trials with low risk of bias on allocation concealment. Only 12 (21.0%) trials had low risk of bias on both blinding of participants and outcome assessment. The risk of bias assessment results are shown in Appendix 6.

Network meta-analysis results

Primary outcome – pregnancy

We performed a network meta-analysis that included 57 randomised controlled trials reporting on 8,082 women. Of these, 19 evaluated a combination of clomiphene and metformin (1,031 women). The remaining trials offered a single treatment in each intervention, including clomiphene (52 trials; 3,511 women), letrozole (21 trials; 1,758 women), metformin alone (14 trials; 910 women), tamoxifen (4 trials; 327 women), follicle-stimulating hormone (2 trials; 197 women), laparoscopic ovarian drilling (1 trial; 36 women) and placebo or no treatment (8 trials; 312 women).

The results of the network meta-analysis are shown in Figure 2 and Table 1. Compared with placebo or no intervention, all the treatment options, except for laparoscopic ovarian drilling, resulted in a significant higher chance of pregnancy. Compared to clomiphene alone, letrozole (odds ratio 1.58, 95% confidence interval 1.25 to 2.00) as well as the combination of clomiphene and metformin (odds ratio 1.81, 95% confidence interval 1.35 to 2.42) led to significant higher pregnancy rates. Similar differences could be found when comparing these two interventions to tamoxifen. The combination of clomiphene and metformin also led to a significant higher pregnancy compared to metformin alone (odds ratio 1.71, 95% confidence interval 1.15 to 2.53).

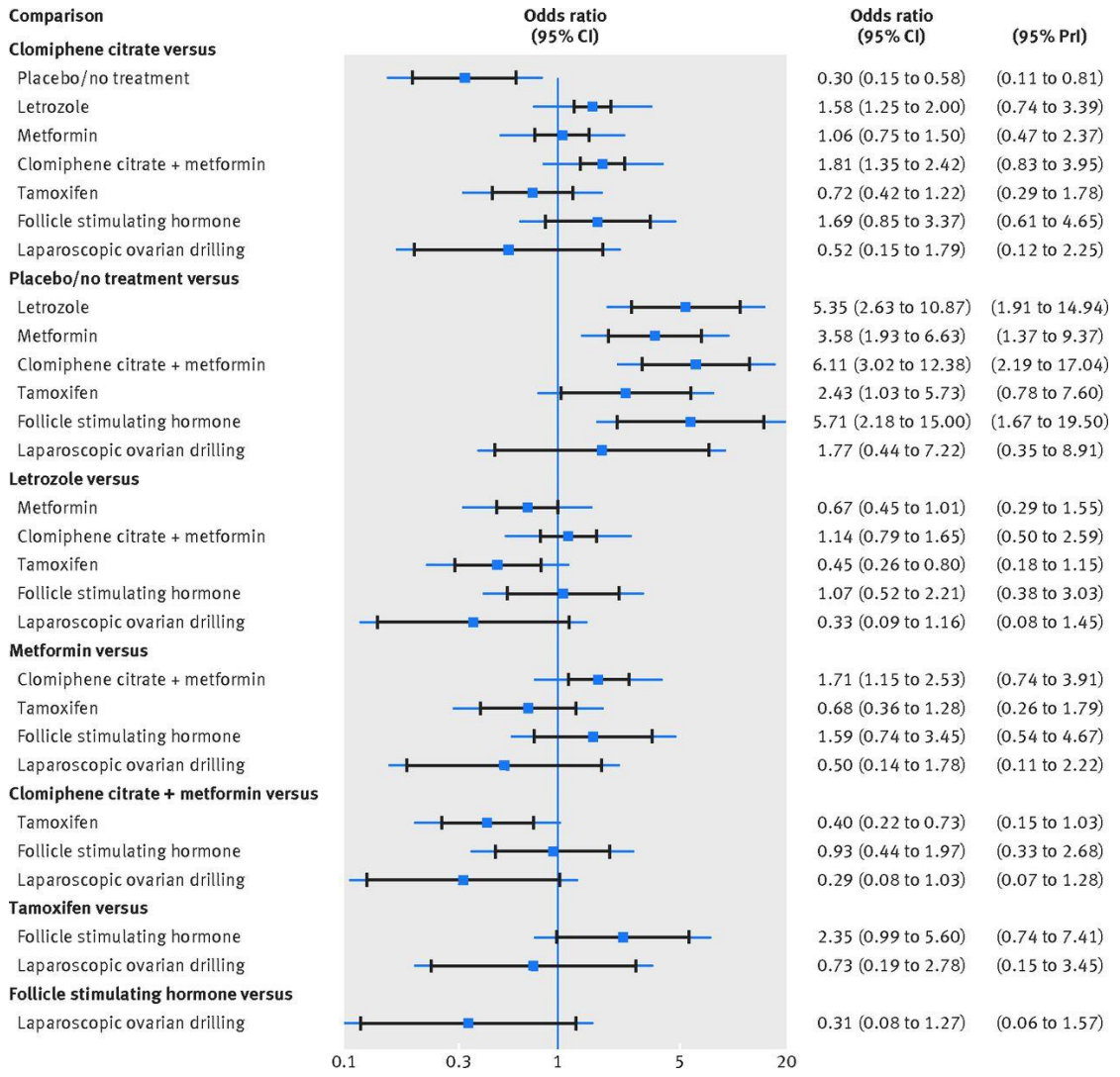


Fig 2 Network meta-analysis of effectiveness of treatment options for pregnancy in women with WHO group II anovulation.

Blue squares=estimate summary odds ratios of each comparison; black horizontal lines=confidence intervals; blue horizontal lines (overall length of lines)=predictive intervals (PrI); blue vertical line=line of no effect (odds ratio=1). Odds ratios less than 1 favour the first intervention; odds ratios greater than 1 favour the second intervention.

When considering predictive intervals in a network meta-analysis, clomiphene, letrozole, metformin, follicle-stimulating hormone and combined clomiphene-metformin still led to higher pregnancy rates compared to placebo or no intervention. For those interventions compared directly, the results from pairwise meta-analysis and network meta-analysis were consistent, apart from follicle-stimulating hormone versus clomiphene (Table 1).

The surface under the cumulative ranking curve was used to provide a hierarchical ranking of the different treatments. The efficacy of every intervention, expressed as a percentage was considered in relation to an imaginary intervention assumed to be the best. Higher surface

under the cumulative ranking curve values therefore correspond to more effective treatments (Chaimani et al., 2013). The surface under the cumulative ranking curve values for the eight ovulation induction regimens were 90%, 82%, 80%, 50%, 46%, 27%, 22% and 3%, for combined clomiphene-metformin, follicle-stimulating hormone, letrozole, metformin, clomiphene, tamoxifen, laparoscopic ovarian drilling and placebo/no treatment, respectively (Appendix 9).

Table 1 Results from pairwise meta-analysis (where possible) and network meta-analysis for primary outcome (pregnancy) in women with WHO group II anovulation

Treatment comparison*	**Pairwise meta-analysis**		**Network meta-analysis**	
	No of studies	Odds ratio (95% CI)	Odds ratio (95% CI)	95% PrI
Clomiphene citrate versus:				
Placebo or no treatment	3	0.20 (0.05 to 0.74)	0.30 (0.15 to 0.58)	0.11 to 0.81
Letrozole	21	1.53 (1.26 to 1.85)	1.58 (1.25 to 2.00)	0.74 to 3.39
Metformin	9	1.10 (0.62 to 1.95)	1.06 (0.75 to 1.50)	0.47 to 2.37
Clomiphene citrate + metformin	19	1.56 (1.24 to 1.97)	1.81 (1.35 to 2.42)	0.83 to 3.95
Tamoxifen	4	0.64 (0.36 to 1.12)	0.72 (0.42 to 1.22)	0.29 to 1.78
Follicle stimulating hormone	2	1.57 (1.04 to 2.37)	1.69 (0.85 to 3.37)	0.61 to 4.65
Laparoscopic ovarian drilling	1	0.52 (0.19 to 1.44)	0.52 (0.15 to 1.79)	0.12 to 2.25
Placebo or no treatment versus:				
Letrozole	NA	NA	5.35 (2.63 to 10.87)	1.91 to 14.94
Metformin	5	3.58 (2.06 to 6.21)	3.58 (1.93 to 6.63)	1.37 to 9.37
Clomiphene citrate + metformin	NA	NA	6.11 (3.02 to 12.38)	2.19 to 17.04
Tamoxifen	NA	NA	2.43 (1.03 to 5.73)	0.78 to 7.60
Follicle stimulating hormone	NA	NA	5.71 (2.18 to 15.00)	1.67 to 19.50
Laparoscopic ovarian drilling	NA	NA	1.77 (0.44 to 7.22)	0.35 to 8.91
Letrozole versus:				
Metformin	1	0.73 (0.41 to 1.32)	0.67 (0.45 to 1.01)	0.29 to 1.55
Clomiphene citrate + metformin	NA	NA	1.14 (0.79 to 1.65)	0.50 to 2.59
Tamoxifen	1	0.67 (0.30 to 1.47)	0.45 (0.26 to 0.80)	0.18 to 1.15
Follicle stimulating hormone	NA	NA	1.07 (0.52 to 2.21)	0.38 to 3.03
Laparoscopic ovarian drilling	NA	NA	0.33 (0.09 to 1.16)	0.08 to 1.45
Metformin versus:				
Clomiphene citrate + metformin	5	1.92 (0.90 to 4.06)	1.71 (1.15 to 2.53)	0.74 to 3.91

Tamoxifen	NA	NA	0.68	(0.36	to	0.26	to
			1.28)			1.79	
Follicle stimulating hormone	NA	NA	1.59	(0.74	to	0.54	to
			3.45)			4.67	
Laparoscopic ovarian drilling	NA	NA	0.50	(0.14	to	0.11	to
			1.78)			2.22	
Clomiphene citrate + metformin versus:							
Tamoxifen	NA	NA	0.40	(0.22	to	0.15	to
			0.73)			1.03	
Follicle stimulating hormone	NA	NA	0.93	(0.44	to	0.33	to
			1.97)			2.68	
Laparoscopic ovarian drilling	NA	NA	0.29	(0.08	to	0.07	to
			1.03)			1.28	
Tamoxifen versus:							
Follicle stimulating hormone	NA	NA	2.35	(0.99	to	0.74	to
			5.60)			7.41	
Laparoscopic ovarian drilling	NA	NA	0.73	(0.19	to	0.15	to
			2.78)			3.45	
Follicle stimulating hormone versus:							
Laparoscopic ovarian drilling	NA	NA	0.31	(0.08	to	0.06	to
			1.27)			1.57	

*Odds ratios less than 1 favour the first intervention; odds ratios greater than 1 favour the second intervention. PrI=predictive interval; NA=not available.

Secondary outcomes

Live birth

For the outcome live birth, 23 randomised controlled trials with 4,206 women were included in the network meta-analysis. Letrozole resulted in a significantly higher live birth rate compared with clomiphene (odds ratio 1.67, 95% confidence interval 1.11 to 2.49) or metformin alone (odds ratio 1.86, 95% confidence interval 1.02 to 3.41). The other comparisons showed no significant differences (Appendix 13).

In terms of live birth, letrozole had the highest surface under the cumulative ranking curve value (81%), followed by follicle-stimulating hormone (74%), combined clomiphene-metformin (71%), tamoxifen (48%) clomiphene (36%) and metformin (30%) while placebo/no treatment (10%) had the lowest surface under the cumulative ranking curve value (Appendix 14).

Ovulation

For the outcome ovulation per woman randomised, 40 randomised controlled trials were included in the network meta-analysis. Compared with placebo, all interventions, except for laparoscopic ovarian drilling, led to a significantly higher ovulation rate. These significances remained similar in the network meta-analysis including predictive intervals.

Letrozole (odds ratio 1.99, 95% confidence interval 1.38 to 2.87) and the combination of clomiphene and metformin (odds ratio 1.55, 95% confidence interval 1.02 to 2.36) led to a higher ovulation rate than clomiphene alone (Appendix 18). The combination of clomiphene and metformin was superior to metformin alone (odds ratio 2.66, 95% confidence interval 1.54 to 4.60), while metformin was inferior to clomiphene alone (odds ratio 0.58, 95% confidence interval 0.37 to 0.93). Both metformin (odds ratio 0.29, 95% confidence interval 0.17 to 0.52) and tamoxifen (odds ratio 0.37, 95% confidence interval 0.16 to 0.81) were inferior to letrozole.

Follicle-stimulating hormone had the highest surface under the cumulative ranking curve value (88%) in terms of ovulation, followed by letrozole (86%), combined clomiphene-metformin (75%), clomiphene (51%), laparoscopic ovarian drilling (39%), tamoxifen (36%), metformin (26%) and placebo/no treatment (1%) (Appendix 19).

Miscarriage

For the outcome miscarriage, after the exclusion of trials with 0 or 100% event rates in all interventions, we included 27 randomised controlled trials in the network meta-analysis. We failed to find any significant difference between each comparison in terms of miscarriage per woman randomised or miscarriage per pregnancy in the network meta-analysis (Appendix 23, 24).

Multiple pregnancy

Twenty trials assessed the outcome multiple pregnancy. When expressed per woman randomized, follicle-stimulating hormone led to higher multiple pregnancy rates than metformin (odds ratio 16.27, 95% confidence interval 1.59 to 166.49). This difference remained significant in network meta-analysis including predictive intervals. Follicle-stimulating hormone also led to higher multiple pregnancy rate than letrozole (odds ratio 7.84, 95% confidence interval 1.10 to 55.90). Both letrozole (odds ratio 0.46, 95% confidence interval 0.23 to 0.92) and metformin (odds ratio 0.22, 95% confidence interval 0.05 to 0.92) led to lower multiple pregnancy rates than clomiphene alone, but these differences were not statistically significant in network meta-analysis including predictive intervals (Appendix 29).

Follicle-stimulating hormone had the highest surface under the cumulative ranking curve value (93%), followed by clomiphene (70%), placebo (50%), tamoxifen (46%), combined clomiphene-metformin (44%), letrozole (34%) and then metformin (14%) (Appendix 30).

Sensitivity analysis results

When the analyses were restricted to studies reporting clinical pregnancy (Appendix 34), the results were consistent with the main findings: letrozole and combination of clomiphene and metformin were superior to clomiphene alone. However, in studies with treatment naïve women or studies with low risk of both randomisation and allocation bias, letrozole remained superior to clomiphene (odds ratio 1.80, 95%confidence interval 1.20 to 2.70; odds ratio 1.97, 95%confidence interval 1.18 to 3.30), while the difference between combined clomiphene-metformin and clomiphene was not statistically significant (odds ratio 1.65, 95% confidence interval 0.98 to 2.80; odds ratio 1.57, 95% confidence interval 0.96 to 2.57) (Appendix 33 and 35).

DISCUSSION

Summary of key findings

Our systematic review and network meta-analysis on ovulation induction in infertile women with WHO group II anovulation has three key findings. First, all pharmacological treatments were more effective than placebo or no intervention in terms of achieving ovulation and pregnancy. Second, the combination of clomiphene and metformin as well as letrozole on its own, were superior to clomiphene in terms of pregnancy and ovulation, and letrozole was superior to clomiphene in terms of live birth. Last, both metformin and letrozole were associated with a lower risk of multiple pregnancy than clomiphene.

Strengths and limitations

To our knowledge this is the first application of network meta-analysis in ovulation induction, analysing all the available data and providing a unique opportunity to rank ovulation induction treatments in a single pooled analysis. We reported all major reproductive outcomes in infertility trials and performed sensitivity analyses in different dimensions including study population and study quality. We made these attempts to guarantee the stability of the results. Another strength of our systematic review was the fact that we did not exclude non-English articles or trials published as abstracts only. These trials are often excluded from other meta-analyses (Misso et al., 2013; Misso et al., 2012; Roque et al., 2015), but in our meta-analysis they contributed 21% (12/57) of the studies and 16% (1321/8082) of the women. We therefore believe that we have included all relevant published randomised controlled trials on ovulation induction in WHO group II anovulation, thus reducing publication bias as much as possible.

Our study also has limitations. First, we only reported reproductive outcomes in our study and were unable to include other relevant outcomes such as side effects which were not reported in many of the primary publications and the reporting strategies varied from study to study. Metformin, for example, is known to generate gastrointestinal side effects (Tang et

al., 2012), but this could not be analysed in our network meta-analysis as it was not systematically reported in all studies. The use of standardized outcomes in studies on ovulation induction would have improved this aspect of our systematic review (Harbin Consensus Conference Workshop Group, 2014; Harbin Consensus Conference Workshop Group et al., 2014; Khan, 2014). Additional discussion on side effects of combined clomiphene-metformin is available in Appendix 36.

Second, we chose pregnancy, defined preferably as clinical pregnancy, as the primary outcome. While the aim of infertile couples is to have a healthy child, we did so as the overall sample size of studies reporting on pregnancy was significantly higher than the sample size of studies reporting on live birth. Studies published in early 2000s or earlier usually followed up participants till pregnancy. In order to make full use of these data and to improve the validity of the transitivity assumption of comparisons among the network, we chose pregnancy as the primary outcome. The conclusions on the effectiveness of a treatment point are often, but not always in women with PCOS (Chen et al., 2016), in the same direction when based either on pregnancy or live birth, while conclusions based on pregnancy as endpoint are more robust as they have more statistical power (Clarke et al., 2010). Ideally, future randomised controlled trials should adhere to the Harbin consensus on outcomes reporting in infertility trials (Harbin Consensus Conference Workshop Group, 2014; Harbin Consensus Conference Workshop Group et al., 2014).

Third, lifestyle intervention was not analysed in this study. Although lifestyle intervention is recommended in many countries as it leads to higher spontaneous ovulation rates (Legro et al., 2015) and natural conceptions rates (Mutsaerts et al., 2016), the role of lifestyle intervention in conjunction to drug treatment is controversial in current evidence. According to a recent Dutch study, lifestyle intervention preceding infertility treatment does not lead to better reproductive outcomes within two years in obese infertile women (Mutsaerts et al., 2016), whilst lifestyle modification with weight loss before ovulation induction improves ovulation and live birth in PCOS in a US study (Legro et al., 2016).

Last, WHO group II anovulation is a heterogeneous condition with a variety of clinical manifestations. Women with different genetic background or metabolic conditions may respond differently on treatment options. The current systematic review only allowed general comparisons among women with WHO group II anovulation. Due to the various reporting strategies, we chose not to perform subgroup analysis, based on characteristics such as body mass index and hyperandrogenaemia status in this network meta-analysis. Apart from the logistic and governance issues associated with data sharing across different countries, asking the original authors to reanalyse the data can be challenging, in view of the substantial time and effort needed to perform secondary analysis. Additionally, there are a number of practical difficulties with post hoc selection of cut-off values for continuous variables like body mass index. If the distribution of participants according to biological cut-off values (25 or 30kg/m^2) are not balanced across groups, the results of subgroup analysis using this cut-off value could be misleading. Individual participant data meta-analysis would be able to address this issue and allow a more personalized strategy for ovulation induction care.

Research implications

Traditionally, the effectiveness of a new treatment option comes from comparisons with placebo or current standard care. To date, there are no trials comparing letrozole and placebo in treatment naïve women. The current network meta-analysis, however, provides insight in this comparison from indirect comparisons and suggests that trials comparing letrozole to placebo are unnecessary and in our opinion even unethical. Evidence on a head-to-head comparison between letrozole and the combination of clomiphene and metformin is lacking. Therefore new trials comparing these two interventions are needed. Additionally, future trials should also compare new treatment options or new combinations to one of these two strategies to enrich the evidence on first-line management of WHO group II anovulation.

Current evidence showed similar miscarriage rates in women with metformin compared to women with other ovulation induction interventions during periconceptional period. Future studies on the use of metformin during pregnancy in women with WHO group II anovulation, including PCOS, can be beneficial.

Individual participant data meta-analysis on this topic is a necessary next step to find target populations for different ovulation induction interventions and therefore to provide evidence for personally targeted infertility care.

Clinical implications and conclusion

In women with WHO group II anovulation including anovulatory PCOS, expectant management is not recommended, as pharmacological ovulation induction significantly improve pregnancy rate (odds ratios between 2.43 and 6.11) compared to placebo no treatment.

Letrozole can be recommended as first-line treatment due to its higher ovulation, pregnancy, and live birth rate as well as lower multiple pregnancy rate, although the reluctance to adapt such new therapy is common in clinical practice (McCartney and Marshall, 2016). The superiority of letrozole over clomiphene was stable in all sensitivity analyses including modifying the criteria of population (treatment naive), reporting strategies (reporting clinical pregnancy) and quality of included studies (low risk of randomisation and allocation bias). Miscarriage is often discussed in the literature especially in women with PCOS, and data in relation to this are controversial (Palomba et al., 2015). In our study, there were no significant differences in miscarriage rates in different comparisons and therefore the superiority of letrozole over clomiphene in terms of live birth does not seem to be related to a decreased miscarriage rate.

Combined clomiphene-metformin can also be recommended as first-line treatment, despite the lack of evidence to improve live birth rates and the instability in sensitivity analyses (Clarke et al., 2010). Of the 19 studies comparing combined clomiphene-metformin to

clomiphene and/or metformin alone, only 7 studies reported live birth. The reduced sample size in the analysis of live birth affected statistical power for this comparison, and could explain the lack of a statistical significant difference between combined clomiphene-metformin and clomiphene alone. The potential higher chances of side effects should also be taken into account in decision making.

Clomiphene alone is not competitive in the network, in terms of effectiveness (pregnancy, live birth, and ovulation) or safety (multiple pregnancy). Gonadotropin, though an effective treatment option, had the greatest probability of leading to multiple pregnancy. It is therefore not recommended to be the first-line treatment in treatment naïve women with WHO group II anovulation.

Despite the promising results shown in this study, neither letrozole nor metformin are approved for the treatment of anovulation in many countries and continue to be used off-label (Usadi and Merriam, 2015; Vitek et al., 2015). The use of letrozole for ovulation induction is explicitly prohibited in many other countries (Birch Petersen et al., 2016; Palomba, 2015), for example Denmark, except if used in approved clinical trials. As shown in Table 2, some guidelines (Balen et al., 2016; Goodman et al., 2015; Legro et al., 2013; National Health and Medical Research Council, 2015) recommended clomiphene citrate or letrozole as first-line treatment, while letrozole was not included in the scope of other guidelines (Conway et al., 2014; Moghetti et al., 2015; National Institute for Health and Care Excellence, 2013; Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2008a, b; Vause et al., 2010) including the NICE guideline in the UK (National Institute for Health and Care Excellence, 2013). Safety concerns about the use of letrozole in infertility were raised in a study presented at the American Society for Reproductive Medicine 2005 annual meeting, that showed a higher risk of locomotor malformations and cardiac anomalies in newborns (Biljan et al., 2005). However, this study was criticized on account of its methodologic limitations, including small sample size of letrozole group and inappropriate choice of control group (Forman et al., 2007). This study has not been

subsequently published as a peer-reviewed paper. According to current evidence (Appendix 39), the use of letrozole in infertility, including PCOS and unexplained infertility, does not increase the risk of congenital anomalies in newborns (Dehbashi et al., 2009; Diamond et al., 2015; Forman et al., 2007; Legro et al., 2014; Ray et al., 2012; Roy et al., 2012; Sharma et al., 2014; Tatsumi et al., 2016; Tulandi et al., 2006; Wu et al., 2016). These results need to be confirmed by future studies. Moreover, there is an urgent need for long-term follow-up data among the offspring of these interventions to confirm the safety of these interventions and help the subsequent guideline development.

In conclusion, in women with WHO group II anovulation, both letrozole and the combination of clomiphene and metformin are superior to clomiphene alone in terms of ovulation and pregnancy. Letrozole is the only therapy showing a statistically significantly higher live birth rate than clomiphene alone.

Table 2 Recommendations on first line ovulation induction from current guidelines and consensus

Guidelines/Consensus	Condition	First-line ovulation induction
WHO guideline, 2016 (Balen et al., 2016)	PCOS	Clomiphene or letrozole
Australian National Health and Medical Research Council (NHMRC) guideline, 2015 updated (National Health and Medical Research Council, 2015)	PCOS	Clomiphene or letrozole
American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and PCOS Society Disease State Clinical Review, 2015 (Goodman et al., 2015)	PCOS	Clomiphene or letrozole
Italian Society of Endocrinology consensus, 2015 (Moghetti et al., 2015)	PCOS	Clomiphene
European Society of Endocrinology position statement, 2014 (Conway et al., 2014)	PCOS	Clomiphene
The Endocrine Society, 2013 (Legro et al., 2013)	PCOS	Clomiphene or letrozole
The National Institute for Health and Care Excellence (NICE) guideline, 2013 (National Institute for Health and Care Excellence, 2013)	WHO II anovulation	Clomiphene, metformin or clomiphene+metformin
Society of Obstetricians and Gynaecologists of Canada guideline, 2010 (Vause et al., 2010)	PCOS	Clomiphene
ESHRE/ASRM consensus, 2008 (Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2008a, Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2008b)	PCOS	Clomiphene

PCOS=polycystic ovary syndrome; ESHRE/ASRM=European Society of Human Reproduction and Embryology/American Society for Reproductive Medicine.

REFERENCES

Abu Hashim H, Al-Inany H, De Vos M, Tournaye H. Three decades after Gjonnaess's laparoscopic ovarian drilling for treatment of PCOS; what do we know? An evidence-based approach. *Arch Gynecol Obstet* 2013;288: 409-422.

Abuelghar WM, Elkady OS, Khamees AA. Clomiphene citrate alone, in combination with metformin or in combination with pioglitazone as first line therapy in induction of ovulation in infertile women with polycystic ovary syndrome, a randomized controlled trial. *Middle East Fertility Society Journal* 2013;18: 135-141.

Amer S, Fakis A, Smith J, Shaw R, Mahran A. Double blind cross-over randomized controlled trial comparing letrozole versus clomiphene citrate for ovulation induction in women with polycystic ovarian syndrome. *Hum Reprod* 2015;30: i96.

Amer SA, Li TC, Metwally M, Emarh M, Ledger WL. Randomized controlled trial comparing laparoscopic ovarian diathermy with clomiphene citrate as a first-line method of ovulation induction in women with polycystic ovary syndrome. *Hum Reprod* 2009;24: 219-225.

Amsterdam ESHRE ASRM-Sponsored 3rd PCOS Consensus Workshop Group. Consensus on women's health aspects of polycystic ovary syndrome (PCOS). *Hum Reprod* 2012;27: 14-24.

Atay V, Cam C, Muhcu M, Cam M, Karateke A. Comparison of letrozole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. *Journal of international medical research* 2006;34: 73-76.

Ayaz A, Alwan Y, Farooq MU. Metformin-clomiphene citrate vs. clomiphene citrate alone: Polycystic ovarian syndrome. *Journal of Human Reproductive Sciences* 2013;6: 15-18.

Aygen EM, Güzel Z, Özgün T, Atakul T, Şahin Y. The use of letrozole for ovulation induction in infertile women with polycystic ovarian syndrome. *Erciyes Tip Dergisi* 2007;29: 195-200.

Badawy A, Abdel Aal I, Abulatta M. Clomiphene citrate or letrozole for ovulation induction in women with polycystic ovarian syndrome: a prospective randomized trial. *Fertility & Sterility* 2009;92: 849-852.

Badawy A, Gibreal A. Clomiphene citrate versus tamoxifen for ovulation induction in women with PCOS: a prospective randomized trial. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2011;159: 151-154.

Balen AH, Morley LC, Misso M, Franks S, Legro RS, Wijeyaratne CN, Stener-Victorin E, Fauser BC, Norman RJ, Teede H. The management of anovulatory infertility in women with polycystic ovary syndrome: an analysis of the evidence to support the development of global WHO guidance. *Hum Reprod Update* 2016;22: 687-708.

Basirat Z, Kashifard M, Amiri MG. Enhanced ovarian folliclular development by Metformin does not correlate with pregnancy rate: A randomized trial. *International Journal of Fertility and Sterility* 2012;6: 31-36.

Bayar U, Basaran M, Kiran S, Coskun A, Gezer S. Use of an aromatase inhibitor in patients with polycystic ovary syndrome: a prospective randomized trial. *Fertility & Sterility* 2006;86: 1447-1451.

Bayram N, van Wely M, Kaaijk EM, Bossuyt PM, van der Veen F. Using an electrocautery strategy or recombinant follicle stimulating hormone to induce ovulation in polycystic ovary syndrome: randomised controlled trial. *BMJ* 2004;328: 192.

Beigi A. Randomized trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in polycystic ovary syndrome. *Human reproduction (Oxford, England)* 2006;21: i129.

Biljan MM, Hemmings R, Brassard N. The Outcome of 150 Babies Following the Treatment With Letrozole or Letrozole and Gonadotropins. *Fertil Steril* 2005;84: S95.

Birch Petersen K, Pedersen NG, Pedersen AT, Lauritsen MP, la Cour Freiesleben N. Mono-ovulation in women with polycystic ovary syndrome: a clinical review on ovulation induction. *Reprod Biomed Online* 2016;32: 563-583.

Boomsma CM, Eijkemans MJ, Hughes EG, Visser GH, Fauser BC, Macklon NS. A meta-analysis of pregnancy outcomes in women with polycystic ovary syndrome. *Hum Reprod Update* 2006;12: 673-683.

Boostanfar R, Jain JK, Mishell DR, Jr., Paulson RJ. A prospective randomized trial comparing clomiphene citrate with tamoxifen citrate for ovulation induction. *Fertility & Sterility* 2001;75: 1024-1026.

Boudhraa K, Jellouli MA, Amri M, Farhat M, Torkhani F, Gara MF. [Indication of metformin in the management of hormonal dysfunction secondary to polycystic ovarian syndrome: prospective comparative study of 63 cases]. *Tunisie medicale* 2010;88: 335-340.

Brown J, Farquhar C, Beck J, Boothroyd C, Hughes E. Clomiphene and anti-oestrogens for ovulation induction in PCOS. *Cochrane Database Syst Rev* 2009: CD002249.

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

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

Chen ZJ, Shi Y, Sun Y, Zhang B, Liang X, Cao Y, Yang J, Liu J, Wei D, Weng N *et al.* Fresh versus Frozen Embryos for Infertility in the Polycystic Ovary Syndrome. *N Engl J Med* 2016;375: 523-533.

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;94: 1647-1651.

Conway G, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Franks S, Gambineri A, Kelestimur F, Macut D, Micic D, Pasquali R *et al.* The polycystic ovary syndrome: a position statement from the European Society of Endocrinology. *Eur J Endocrinol* 2014;171: P1-29.

Cudmore DW, Tupper WR. Induction of ovulation with clomiphene citrate. A double-blind study. *Fertil Steril* 1966;17: 363-373.

Dasari P, Pranahita GK. The efficacy of metformin and clomiphene citrate combination compared with clomiphene citrate alone for ovulation induction in infertile patients with PCOS. *Journal of Human Reproductive Sciences* 2009;2: 18-22.

Dehbashi S, Dehbashi S, Kazerooni T, Robati M, Alborzi S, Parsanezhad ME, Shadman A. Comparison of the effects of letrozole and clomiphene citrate on ovulation and pregnancy rate in patients with polycystic ovary syndrome. *Iranian Journal of Medical Sciences* 2009;34: 23-28.

Diamond MP, Legro RS, Coutifaris C, Alvero R, Robinson RD, Casson P, Christman GM, Ager J, Huang H, Hansen KR *et al.* Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility. *N Engl J Med* 2015;373: 1230-1240.

El-Biely MM, Habba M. The use of metformin to augment the induction of ovulation in obese infertile patients with polycystic ovary syndrome. *Middle East Fertility Society Journal* 2001;6: 43-49.

ESHRE Capri Workshop Group. Health and fertility in World Health Organization group 2 anovulatory women. *Hum Reprod Update* 2012;18: 586-599.

Farquhar C, Brown J, Marjoribanks J. Laparoscopic drilling by diathermy or laser for ovulation induction in anovulatory polycystic ovary syndrome. *Cochrane Database Syst Rev* 2012;6: CD001122.

Fauser BC, Tarlatzis BC, Rebar RW, Legro RS, Balen AH, Lobo R, Carmina E, Chang J, Yildiz BO, Laven JS *et al.* Consensus on women's health aspects of polycystic ovary syndrome (PCOS): the Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. *Fertil Steril* 2012;97: 28-38 e25.

Fleming R, Hopkinson ZE, Wallace AM, Greer IA, Sattar N. Ovarian function and metabolic factors in women with oligomenorrhea treated with metformin in a randomized double blind placebo-controlled trial. *Journal of Clinical Endocrinology & Metabolism* 2002;87: 569-574.

Forman R, Gill S, Moretti M, Tulandi T, Koren G, Casper R. Fetal safety of letrozole and clomiphene citrate for ovulation induction. *J Obstet Gynaecol Can* 2007;29: 668-671.

Franik S, Kremer JA, Nelen WL, Farquhar C. Aromatase inhibitors for subfertile women with polycystic ovary syndrome. *Cochrane Database Syst Rev* 2014;2: CD010287.

Garcia CR, Freeman EW, Rickels K, Wu C, Scholl G, Galle PC, Boxer AS. Behavioral and emotional factors and treatment responses in a study of anovulatory infertile women. *Fertility & Sterility* 1985;44: 478-483.

Goodman NF, Cobin RH, Futterweit W, Glueck JS, Legro RS, Carmina E, American Association of Clinical E, American College of E, Androgen E, Society P. American Association of Clinical Endocrinologists, American College of Endocrinology, and Androgen Excess and Pcos Society Disease State Clinical Review: Guide to the Best Practices in the Evaluation and Treatment of Polycystic Ovary Syndrome - Part 2. *Endocr Pract* 2015;21: 1415-1426.

Hansen M, Kurinczuk JJ, Milne E, de Klerk N, Bower C. Assisted reproductive technology and birth defects: a systematic review and meta-analysis. *Hum Reprod Update* 2013;19: 330-353.

Harbin Consensus Conference Workshop Group. Improving the Reporting of Clinical Trials of Infertility Treatments (IMPRINT): modifying the CONSORT statement. *Fertil Steril* 2014;102: 952-959 e915.

Harbin Consensus Conference Workshop Group, Conference C, Legro RS, Wu X, Scientific C, Barnhart KT, Farquhar C, Fauser BC, Mol B. Improving the reporting of clinical trials of infertility treatments (IMPRINT): modifying the CONSORT statementdaggerdouble dagger. *Hum Reprod* 2014;29: 2075-2082.

Hart R, Norman RJ. The longer-term health outcomes for children born as a result of IVF treatment: Part I--General health outcomes. *Hum Reprod Update* 2013;19: 232-243.

Higgins JPT, Green, S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. The Cochrane Collaboration.

Homburg R, Hendriks ML, Konig TE, Anderson RA, Balen AH, Brincat M, Child T, Davies M, D'Hooghe T, Martinez A *et al.* Clomifene citrate or low-dose FSH for the first-line

treatment of infertile women with anovulation associated with polycystic ovary syndrome: a prospective randomized multinational study. *Hum Reprod* 2012;27: 468-473.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162: 777-784.

Jahan S. Comparative study of efficacy among metformin, clomiphene citrate and aromatase inhibitor (letrozole) as the first-line medication for ovulation induction, achievement of pregnancy and live birth in Asian women with polycystic ovarian syndrome: A prospective trial. *International Journal of Gynecology and Obstetrics* 2015;131: E503.

Jansen JP, Fleurence R, Devine B, Itzler R, Barrett A, Hawkins N, Lee K, Boersma C, Annemans L, Cappelleri JC. Interpreting indirect treatment comparisons and network meta-analysis for health-care decision making: report of the ISPOR Task Force on Indirect Treatment Comparisons Good Research Practices: part 1. *Value Health* 2011;14: 417-428.

Johnson JE, Jr., Cohen MR, Goldfarb AF, Rakoff AE, Kistner RW, Plotz EJ, Vorys N. The efficacy of clomiphene citrate for induction of ovulation. A controlled study. *Int J Fertil* 1966;11: 265-270.

Johnson NP, Stewart AW, Falkiner J, Farquhar CM, Milsom S, Singh VP, Okonkwo QL, Buckingham KL, React-Nz am-cftg. PCOSMIC: a multi-centre randomized trial in women with PolyCystic Ovary Syndrome evaluating Metformin for Infertility with Clomiphene. *Hum Reprod* 2010;25: 1675-1683.

Kar S. Clomiphene citrate or letrozole as first-line ovulation induction drug in infertile PCOS women: A prospective randomized trial. *Journal of Human Reproductive Sciences* 2012;5: 262-265.

Kar S, Sanchita S. Clomiphene citrate, metformin or a combination of both as the first line ovulation induction drug for Asian Indian women with polycystic ovarian syndrome: A randomized controlled trial. *Journal of Human Reproductive Sciences* 2015;8: 197-201.

Karimzadeh MA, Eftekhar M, Taheripanah R, Tayebi N, Sakhavat L, Zare F. The effect of administration of metformin on lipid profile changes and insulin resistance in patients with polycystic ovary syndrome. *Middle East Fertility Society Journal* 2007;12: 174-178.

Karimzadeh MA, Javedani M. An assessment of lifestyle modification versus medical treatment with clomiphene citrate, metformin, and clomiphene citrate-metformin in patients with polycystic ovary syndrome. *Fertility & Sterility* 2010;94: 216-220.

Keikha F, Shahraki Mojahed B. Induction ovulation in polycystic ovary patient with clomiphene citrate and letrozole. *Iranian Journal of Reproductive Medicine* 2011;9: 46.

Khan K. The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *BJOG* 2014;121: 1181-1182.

Khorram O, Helliwell JP, Katz S, Bonpane CM, Jaramillo L. Two weeks of metformin improves clomiphene citrate-induced ovulation and metabolic profiles in women with polycystic ovary syndrome. *Fertility & Sterility* 2006;85: 1448-1451.

Leanza V, Coco L, Grasso F, Leanza G, Zarbo G, Palumbo M. Ovulation induction with clomiphene citrate and metformin in women with polycystic ovary syndrome. *Minerva ginecologica* 2014;66: 299-301.

Legro RS, Arslanian SA, Ehrmann DA, Hoeger KM, Murad MH, Pasquali R, Welt CK, Endocrine S. Diagnosis and treatment of polycystic ovary syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2013;98: 4565-4592.

Legro RS, Barnhart HX, Schlaff WD, Carr BR, Diamond MP, Carson SA, Steinkampf MP, Coutifaris C, McGovern PG, Cataldo NA *et al.* Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. *New England journal of medicine* 2007;356: 551-566.

Legro RS, Brzyski RG, Diamond MP, Coutifaris C, Schlaff WD, Casson P, Christman GM, Huang H, Yan Q, Alvero R *et al.* Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. *N Engl J Med* 2014;371: 119-129.

Legro RS, Dodson WC, Kris-Etherton PM, Kunselman AR, Stetter CM, Williams NI, Gnatuk CL, Estes SJ, Fleming J, Allison KC *et al.* Randomized Controlled Trial of Preconception Interventions in Infertile Women With Polycystic Ovary Syndrome. *J Clin Endocrinol Metab* 2015;100: 4048-4058.

Legro RS, Dodson WC, Kunselman AR, Stetter CM, Kris-Etherton PM, Williams NI, Gnatuk CL, Estes SJ, Allison KC, Sarwer DB *et al.* Benefit of Delayed Fertility Therapy with Preconception Weight Loss over Immediate Therapy in Obese Women with PCOS. *J Clin Endocrinol Metab* 2016: jc20161659.

Liu C, Feng G, Wang Q, Xiao L, Huang W. Comparison of ovulation induction protocol for women with polycystic ovarian syndrome: A prospective randomized trial. *International Journal of Gynecology and Obstetrics* 2015;131: E231-E232.

Lopez E, Gunby J, Daya S, Parrilla JJ, Abad L, Balasch J. Ovulation induction in women with polycystic ovary syndrome: randomized trial of clomiphene citrate versus low-dose recombinant FSH as first line therapy. *Reproductive Biomedicine Online* 2004;9: 382-390.

Lord J, Thomas R, Fox B, Acharya U, Wilkin T. The effect of metformin on fat distribution and the metabolic syndrome in women with polycystic ovary syndrome--a randomised, double-blind, placebo-controlled trial. *BJOG: An International Journal of Obstetrics & Gynaecology* 2006;113: 817-824.

Lorzadeh N, Kazemirad S, Mohammadi Z. Comparison of effects letrozole and clomiphene citrate for ovulation induction in women with polycystic ovary syndrome. *Iranian Journal of Obstetrics, Gynecology and Infertility* 2011;14.

Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Stat Med* 2004;23: 3105-3124.

Maged AM, Elsawah H, Abdelhafez A, Bakry A, Mostafa WAI. The adjuvant effect of metformin and N-acetylcysteine to clomiphene citrate in induction of ovulation in patients with Polycystic Ovary Syndrome. *Gynecological endocrinology* 2015.

McCartney CR, Marshall JC. CLINICAL PRACTICE. Polycystic Ovary Syndrome. *N Engl J Med* 2016;375: 54-64.

Mills EJ, Thorlund K, Ioannidis JP. Demystifying trial networks and network meta-analysis. *BMJ* 2013;346: f2914.

Misso ML, Costello MF, Garrubba M, Wong J, Hart R, Rombauts L, Melder AM, Norman RJ, Teede HJ. Metformin versus clomiphene citrate for infertility in non-obese women with polycystic ovary syndrome: a systematic review and meta-analysis. *Hum Reprod Update* 2013;19: 2-11.

Misso ML, Wong JL, Teede HJ, Hart R, Rombauts L, Melder AM, Norman RJ, Costello MF. Aromatase inhibitors for PCOS: a systematic review and meta-analysis. *Hum Reprod Update* 2012;18: 301-312.

Mobusher I. Comparison of the efficacy of letrozole and clomiphene citrate for ovulation induction in infertile women with polycystic ovary syndrome. *Pakistan Journal of Medical and Health Sciences* 2014;8: 905-908.

Moghetti P, Carmina E, De Leo V, Lanzone A, Orio F, Pasquali R, Toscano V. How to manage the reproductive issues of PCOS: a 2015 integrated endocrinological and gynecological consensus statement of the Italian Society of Endocrinology. *J Endocrinol Invest* 2015;38: 1025-1037.

Moll E, Bossuyt PM, Korevaar JC, Lambalk CB, van der Veen F. Effect of clomifene citrate plus metformin and clomifene citrate plus placebo on induction of ovulation in women with newly diagnosed polycystic ovary syndrome: randomised double blind clinical trial. *BMJ (Clinical research ed)* 2006;332: 1485.

Moll E, van der Veen F, van Wely M. The role of metformin in polycystic ovary syndrome: a systematic review. *Hum Reprod Update* 2007;13: 527-537.

Mutsaerts MA, van Oers AM, Groen H, Burggraaff JM, Kuchenbecker WK, Perquin DA, Koks CA, van Golde R, Kaaijk EM, Schierbeek JM *et al.* Randomized Trial of a Lifestyle Program in Obese Infertile Women. *N Engl J Med* 2016;374: 1942-1953.

Nahuis MJ, Kose N, Bayram N, van Dessel HJ, Braat DD, Hamilton CJ, Hompes PG, Bossuyt PM, Mol BW, van der Veen F *et al.* Long-term outcomes in women with polycystic ovary syndrome initially randomized to receive laparoscopic electrocautery of the ovaries or ovulation induction with gonadotrophins. *Hum Reprod* 2011;26: 1899-1904.

Nahuis MJ, Oude Lohuis E, Kose N, Bayram N, Hompes P, Oosterhuis GJ, Kaaijk EM, Cohlen BJ, Bossuyt PP, van der Veen F *et al.* Long-term follow-up of laparoscopic electrocautery of the ovaries versus ovulation induction with recombinant FSH in clomiphene citrate-resistant women with polycystic ovary syndrome: an economic evaluation. *Hum Reprod* 2012;27: 3577-3582.

Nahuis MJ, Oude Lohuis EJ, Bayram N, Hompes PG, Oosterhuis GJ, van der Veen F, Mol BW, van Wely M. Pregnancy complications and metabolic disease in women with clomiphene citrate-resistant anovulation randomized to receive laparoscopic electrocautery of the ovaries or ovulation induction with gonadotropins: a 10-year follow-up. *Fertil Steril* 2014;101: 270-274.

National Health and Medical Research Council. Evidence-based guideline for the assessment and management of polycystic ovary syndrome (Updated August 2015 - Section 7.4 Aromatase inhibitors). *Australian Government NHMRC Website* 2015.

National Institute for Health and Care Excellence. Fertility: Assessment and Treatment for People with Fertility Problems *National Institute for Health and Clinical Excellence: Guidance*. 2013, London.

Nazik H, Kumtepe Y. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction for women with polycystic ovarian syndrome. *HealthMED* 2012;6: 879-883.

Palomba S. Aromatase inhibitors for ovulation induction. *J Clin Endocrinol Metab* 2015;100: 1742-1747.

Palomba S, de Wilde MA, Falbo A, Koster MP, La Sala GB, Fauser BC. Pregnancy complications in women with polycystic ovary syndrome. *Hum Reprod Update* 2015;21: 575-592.

Palomba S, Orio F, Jr., Falbo A, Manguso F, Russo T, Cascella T, Tolino A, Carmina E, Colao A, Zullo F. Prospective parallel randomized, double-blind, double-dummy controlled clinical trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in nonobese anovulatory women with polycystic ovary syndrome. *Journal of Clinical Endocrinology & Metabolism* 2005;90: 4068-4074.

Pandey S, Shetty A, Hamilton M, Bhattacharya S, Maheshwari A. Obstetric and perinatal outcomes in singleton pregnancies resulting from IVF/ICSI: a systematic review and meta-analysis. *Hum Reprod Update* 2012;18: 485-503.

Pinborg A, Wennerholm UB, Romundstad LB, Loft A, Aittomaki K, Soderstrom-Anttila V, Nygren KG, Hazekamp J, Bergh C. Why do singletons conceived after assisted reproduction technology have adverse perinatal outcome? Systematic review and meta-analysis. *Hum Reprod Update* 2013;19: 87-104.

Raja A, Hashmi SN, Sultana N, Rashid H. Presentation of polycystic ovary syndrome and its management with clomiphene alone and in combination with metformin. *J Ayub Med Coll Abbottabad* 2005;17: 50-53.

Ray PB, Ray A, Chakraborti PS. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction in Indian women with polycystic ovarian syndrome. *Archives of Gynecology & Obstetrics* 2012;285: 873-877.

Robinson R, Swezey M, Propst A, Bates G. Metformin added to clomiphene citrate does not improve pregnancy rates in hyperandrogenic, chronic anovulatory women: A randomized trial. *Fertil Steril* 2003;80: S273-S274, Abstract no: P.

Roque M, Tostes AC, Valle M, Sampaio M, Geber S. Letrozole versus clomiphene citrate in polycystic ovary syndrome: systematic review and meta-analysis. *Gynecol Endocrinol* 2015;31: 917-921.

Rotterdam ESHRE ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004a;81: 19-25.

Rotterdam ESHRE ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004b;19: 41-47.

Roy K, Baruah J, Singla S, Sharma J, Singh N, Jain S, Goyal M. A prospective randomized trial comparing the efficacy of Letrozole and Clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. *Journal of Human Reproductive Sciences* 2012;5: 20-25.

Sahin Y, Yirmibes U, Kelestimur F, Aygen E. The effects of metformin on insulin resistance, clomiphene-induced ovulation and pregnancy rates in women with polycystic ovary syndrome. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2004;113: 214-220.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods* 2012;3: 80-97.

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

Santonocito V, Rapisarda V, Abruzzo SRM, Pollicino R, Coco L, Zarbo G. Comparison between clomiphene citrate and metformin for induction of ovulatory cycles in infertile nonobese women with polycystic ovary syndrome. *Giornale Italiano di Ostetricia e Ginecologia* 2009;31: 455-460.

Selim MF, Borg TF. Letrozole and clomiphene citrate effect on endometrial and subendometrial vascularity in treating infertility in women with polycystic ovary syndrome. 2012, pp. 405-410.

Seyedoshohadaei F, Zandvakily F, Shahgeibi S. Comparison of the effectiveness of clomiphene citrate, tamoxifen and letrozole in ovulation induction in infertility due to isolated unovulation. *Iranian Journal of Reproductive Medicine* 2012;10: 531-536.

Sharief M, Nafee NR. Comparison of letrazole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. *Journal of the Pakistan Medical Association* 2015;65: 1149-1152.

Sharma S, Ghosh S, Singh S, Chakravarty A, Ganesh A, Rajani S, Chakravarty BN. Congenital malformations among babies born following letrozole or clomiphene for infertility treatment. *PLoS One* 2014;9: e108219.

Sheikh-El-Arab Elsedeek M, Elmaghraby HAH. Predictors and characteristics of letrozole induced ovulation in comparison with clomiphene induced ovulation in anovulatory PCOS women. *Middle East Fertility Society Journal* 2011;16: 125-130.

Sirmans SM, Pate KA. Epidemiology, diagnosis, and management of polycystic ovary syndrome. *Clin Epidemiol* 2013;6: 1-13.

Song F, Loke YK, Walsh T, Glenny AM, Eastwood AJ, Altman DG. Methodological problems in the use of indirect comparisons for evaluating healthcare interventions: survey of published systematic reviews. *BMJ* 2009;338: b1147.

Stein IF, Leventhal ML. Amenorrhea associated with bilateral polycystic ovaries. *Am J Obstet Gynecol* 1935;29: 181-191.

Tang T, Glanville J, Hayden CJ, White D, Barth JH, Balen AH. Combined lifestyle modification and metformin in obese patients with polycystic ovary syndrome. A randomized, placebo-controlled, double-blind multicentre study. *Hum Reprod* 2006;21: 80-89.

Tang T, Lord JM, Norman RJ, Yasmin E, Balen AH. Insulin-sensitising drugs (metformin, rosiglitazone, pioglitazone, D-chiro-inositol) for women with polycystic ovary syndrome, oligo amenorrhoea and subfertility. *Cochrane Database Syst Rev* 2012;5: CD003053.

Tatsumi T, Jwa SC, Kuwahara A, Irahara M, Kubota T, Saito H. No increased risk of major congenital anomalies or adverse pregnancy or neonatal outcomes following letrozole use in assisted reproductive technology. *Hum Reprod* 2016.

Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Consensus on infertility treatment related to polycystic ovary syndrome. *Hum Reprod* 2008a;23: 462-477.

Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Consensus on infertility treatment related to polycystic ovary syndrome. *Fertil Steril* 2008b;89: 505-522.

Tulandi T, Martin J, Al-Fadhli R, Kabli N, Forman R, Hitkari J, Librach C, Greenblatt E, Casper RF. Congenital malformations among 911 newborns conceived after infertility treatment with letrozole or clomiphene citrate. *Fertil Steril* 2006;85: 1761-1765.

Usadi RS, Merriam KS. On-label and off-label drug use in the treatment of female infertility. *Fertil Steril* 2015;103: 583-594.

Vause TD, Cheung AP, Sierra S, Claman P, Graham J, Guillemin JA, Lapensee L, Stewart S, Wong BC, Society of O *et al.* Ovulation induction in polycystic ovary syndrome. *J Obstet Gynaecol Can* 2010;32: 495-502.

Vegetti W, Riccaboni A, Columbo M, Baroni E, Diaferia D, Ragni G, Crosignani PG. Randomized study of induction of ovulation by two different molecules with antioestrogenic effects, in patients with chronic anovulation disorders. 1999;72: S234-235.

Vitek W, Alur S, Hoeger KM. Off-label drug use in the treatment of polycystic ovary syndrome. *Fertil Steril* 2015;103: 605-611.

Williams CD, Pastore LM, Shelly WB, Bailey AP, Baras DC, Bateman BG. A randomized, placebo-controlled study of the influence of instant-release metformin on response to clomiphene citrate and time to conception in polycystic ovary syndrome. *Fertil Steril* 2009;92: S105.

Wu XK, Wang YY, Liu JP, Liang RN, Xue HY, Ma HX, Shao XG, Ng EH, Reproductive, Developmental Network in Chinese M. Randomized controlled trial of letrozole, berberine,

or a combination for infertility in the polycystic ovary syndrome. *Fertil Steril* 2016;106: 757-765 e751.

Zain MM, Jamaluddin R, Ibrahim A, Norman RJ. Comparison of clomiphene citrate, metformin, or the combination of both for first-line ovulation induction, achievement of pregnancy, and live birth in Asian women with polycystic ovary syndrome: a randomized controlled trial. *Fertility & Sterility* 2009;91: 514-521.

Zeinalzadeh M, Basirat Z, Esmailpour M. Efficacy of letrozole in ovulation induction compared to that of clomiphene citrate in patients with polycystic ovarian syndrome. *Journal of reproductive medicine* 2010;55: 36-40.

Chapter 4

Ovulation induction for women with polycystic ovary syndrome treated unsuccessfully with clomiphene citrate: systematic review and network meta-analysis

Wang R, Jones ER, Costello MF, Bhattacharya S, Legro RS, Ng E, Johnson NP, Norman RJ, van Wely M, Mol BWJ. Ovulation induction for women with polycystic ovary syndrome treated unsuccessfully with clomiphene citrate: systematic review and network meta-analysis. 2019 unsubmitted.

Statement of Authorship

Title of Paper	Ovulation induction for women with polycystic ovary syndrome treated unsuccessfully with clomiphene citrate: systematic review and network meta-analysis
Publication Status	☐ Published ☐ Accepted for Publication ☐ Submitted for Publication ☑ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Wang R, Jones ER, Costello MF, Bhattacharya S, Legro RS, Ng E, Johnson NP, Norman RJ, van Wely M, Mol BWJ. Ovulation induction for women with polycystic ovary syndrome treated unsuccessfully with clomiphene citrate: systematic review and network meta-analysis. 2019 unpublished.

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	RW contributed to the study conception and design, collected the data, interpreted the work, drafted the manuscript, commented on the drafts, and approved the final draft.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	10/9/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Esther R Jones		
Contribution to the Paper	ERJ collected the data, interpreted the work, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Michael F Costello
Contribution to the Paper	MFC contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.

Signature		Date	10/9/2019

Name of Co-Author	Siladitya Bhattacharya		
Contribution to the Paper	SB contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Richard S Legro		
Contribution to the Paper	RSL contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Ernest Hung Yu Ng		
Contribution to the Paper	RSL contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Neil P Johnson		
Contribution to the Paper	NPJ contributed to the study conception, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Robert J Norman		
Contribution to the Paper	NPJ contributed to the study conception, collected the data, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Madelon van Wely

Contribution to the Paper	MVW contributed to the study conception, collected the data, interpreted the work, critically revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

Name of Co-Author	Ben Willem J Mol		
Contribution to the Paper	BWJM contributed to the study conception, collected the data, interpreted the work, drafted the manuscript, commented on the drafts and approved the final draft.		
Signature		Date	10/9/2019

ABSTRACT

Objective: To compare the effectiveness of different ovulation induction options, including clomiphene citrate (CC), letrozole, metformin, gonadotropins, their combinations, or laparoscopic ovarian drilling in women with polycystic ovary syndrome (PCOS) who were treated unsuccessfully with CC, and to identify the best strategy for second-line treatment.

Design: Systematic review and network meta-analysis of randomised controlled trials (RCTs).

Setting: Not applicable.

Patient(s): Women with CC-failure or CC-resistant PCOS.

Intervention(s): CC, letrozole, metformin, gonadotrophins, a combination of these interventions, or laparoscopic ovarian drilling (LOD).

Main outcome measure(s): live birth/ongoing pregnancy and multiple pregnancy.

Results(s): We included 44 RCTs (7260 couples) in this systematic review and 42 RCTs (6925 couples) in a subsequent network meta-analysis. Overall, the certainty of evidence was low to moderate: the main limitations were imprecision and/or risk of bias. Twenty-one RCTs reported data on live birth/ongoing pregnancy in 3,137 women. Compared to CC, letrozole alone and gonadotrophins alone resulted in higher odds of live birth/ongoing pregnancy (OR 1.67, 95% CI 1.16-2.40; and OR 1.63, 95% CI 1.24-2.14 respectively), while the addition of metformin to these two interventions improved live birth/ongoing pregnancy further (OR 4.89, 95% CI 1.10-21.81; and OR 4.47, 95% CI 2.37-8.41 respectively). Gonadotrophins plus metformin, letrozole plus metformin, CC plus gonadotrophins, letrozole and gonadotrophins were the top-ranking interventions, with the surface under the cumulative ranking (SUCRA) values of 92.2%, 88.2%, 58.7% and 57.1%, respectively. Twenty RCTs (3944 women) were included in the network meta-analysis for multiple pregnancy. LOD was the only intervention resulting in fewer multiple pregnancies than CC (OR 0.22, 95% CI 0.07-0.67), while there was insufficient evidence of differences between the other interventions and CC.

Conclusion(s): In women with PCOS treated unsuccessfully with CC, moderate-certainty evidence showed that letrozole or gonadotrophins alone, or in combination with metformin, are the top-ranking interventions for live birth/ongoing pregnancy and therefore these interventions should be considered as the second-line ovulation induction options.

PROSPERO registration number: CRD42017068425

Key words: polycystic ovary syndrome, ovulation induction, clomiphene citrate resistant, clomiphene citrate failure, network meta-analysis.

INTRODUCTION

Polycystic ovary syndrome (PCOS) is the most common cause of women with anovulatory infertility (Balen et al., 2016). Clomiphene citrate (CC) has been conventionally used as the first-line treatment for ovulation induction in women with PCOS related infertility (Balen et al., 2016). Where ovulation and/or conception fails to occur (i.e. CC-resistance and CC-failure), other medical or surgical options are considered before in-vitro fertilisation (IVF) (Balen et al., 2016; Teede et al., 2018). The alternative medical ovulation induction options include letrozole, gonadotrophins, metformin, or a combination of these interventions (Teede et al., 2018). In addition, laparoscopic ovarian drilling (LOD) is a surgical alternative strategy (Teede et al., 2018).

A number of randomised trials (RCTs) have compared these medical and surgical ovulation induction strategies in women in whom CC was unsuccessful and subsequent meta-analyses evaluated these head-to-head comparisons (Abu Hashim et al., 2015; Bordewijk et al., 2017; Farquhar et al., 2012; Franik et al., 2018; Morley et al., 2017; Weiss et al., 2019). Given that there are multiple interventions of interest, it is difficult to find the most effective treatment among these interventions based on the conclusion of pairwise meta-analyses.

An ideal RCT should compare all these available interventions; however, conducting such a trial is never feasible. Network meta-analysis is a potentially useful tool to compare these different interventions and guide clinical practice by incorporating both direct and indirect evidence and ranking multiple treatments based on their summary results (Riley et al., 2017). A previous network meta-analyses had evaluated different interventions for women with CC-resistant PCOS (Yu et al., 2017). However, this network meta-analysis excluded women with CC-failure, missed a number of relevant large trials and was not based on a prespecified protocol. These concerns limit its clinical implication.

We therefore performed a systematic review and network meta-analysis to compare the effectiveness of different treatment options, including CC, letrozole, metformin,

gonadotropins, their combinations and LOD, in women with PCOS who were treated unsuccessfully with CC, and to identify the best strategy for second-line treatment.

MATERIALS AND METHODS

Protocol and search strategies

The protocol of this systematic review was registered on PROSPERO (CRD42017068425). We reported the systematic review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) extension statement for network meta-analysis (Hutton et al., 2015).

We searched MEDLINE, EMBASE, and Cochrane Central Register of Controlled Trials (CENTRAL) to identify eligible trials. Additionally, we searched trial registries including the World Health Organization International Clinical Trials Registry Platform, ClinicalTrials.gov and International Clinical Trials Registry Platform. We also reviewed the reference lists of relevant papers and corresponded with trialists to identify other relevant trials. The key search terms included "PCOS", "clomiphene", and "RCTs". The last search was conducted in 11th July 2019. The detailed search strategies are presented in Supplemental Table 1. We did not apply language restrictions; and included both full publications and abstracts.

Eligibility criteria

We included RCTs comparing any of the following interventions to each other: CC, letrozole, metformin, gonadotropins, laparoscopic ovarian drilling, or a combination of these interventions in women with PCOS treated unsuccessfully with CC.

Studies comparing different doses of the same intervention or different types of gonadotropins were excluded. The primary effectiveness outcome was live birth or ongoing pregnancy. Ongoing pregnancy was only used when live birth was not reported. The primary safety outcome was multiple pregnancy. The secondary outcomes included clinical pregnancy, miscarriage, ovulation and adverse events (ovarian hyperstimulation syndrome [OHSS]).

Study selection and data collection

Two reviewers (from RW, ERJ, BWM) independently evaluated study eligibility and disagreements were solved by discussions with a third reviewer (MvW or RJN). Two reviewers (RW and ERJ) used a predesigned form to collect the following information: name of the first author, publication year, country, funding, study population, participants' baseline characteristics, funding, details of interventions, number of ovulation induction cycles, sample sizes and outcome data.

Quality assessment within individual studies and across studies

Two reviewers (RW and ERJ) evaluated the risks of bias of individual studies in the following domains described in the Cochrane Collaboration Handbook: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective reporting, and other sources of bias (Higgins and Green, 2011). In case of disagreement, a third reviewer (BWM) was involved to reach consensus.

We used a web application, Confidence in Network Meta-analysis (CINeMA), to assess confidence in the results from a network meta-analysis in six domains: within-study bias, across-studies bias, indirectness, imprecision, heterogeneity and incoherence and graded the overall certainty of evidence as high, moderate, low or very low for each comparison (Nikolakopoulou et al., 2019; Salanti et al., 2014).

Statistical analysis

For network meta-analysis, we first used network plots to illustrate the geometry of the network for each outcome (Chaimani and Salanti, 2015). We then assessed the global inconsistency by using design-by-treatment interaction model (White, 2015) and evaluated the local inconsistency by using node-splitting method (Dias et al., 2010). When there was no significant inconsistency, we performed network meta-analyses within multivariate random-effects meta-analysis models and assumed a common heterogeneity variance (White, 2015). We presented odds ratios (ORs) with their 95% confidence intervals (CIs)

and used CC as the reference arm. Studies with zero event in both arms were excluded from the analyses. When both global and local inconsistency were observed, we explored the inconsistent comparisons and adjusted the network. Finally, we used the surface under the cumulative ranking (SUCRA) curve to rank the treatments (Salanti et al., 2011). When direct comparisons are available, we performed pairwise meta-analyses in random-effects model and used I-squared statistic to present heterogeneity (Higgins and Green, 2011). We intended to apply the comparison adjusted funnel plot by using CC as the reference comparison to assess small study effects if the number of included studies were sufficient (Chaimani and Salanti, 2015).

We performed subgroup analysis for studies on women with CC resistant (no ovulation after CC) and CC failure (no pregnancy after CC). We restricted the inclusion to those trials on low risk of randomisation and allocation bias in the sensitivity analysis. We performed all statistical analysis in Stata software (version 15.1, StataCorp LLC) (Chaimani and Salanti, 2015; White, 2015). The unit for all analyses was per women randomised.

RESULTS

Characteristics of included studies

Of the 2619 studies identified, 72 studies were further assessed in full text. Finally, 44 RCTs reporting on 7260 women with PCOS treated unsuccessfully with CC were included (Figure 1) (Abd Elgafor, 2013; Abdellah, 2011; Abu Hashim et al., 2011a; Abu Hashim et al., 2010a; Abu Hashim et al., 2010b; Abu Hashim et al., 2011b; Abu Hashim et al., 2011c; Bayram et al., 2004; Begum et al., 2013; Begum et al., 2009; Davar et al., 2011; De Leo et al., 1999; Farquhar et al., 2002; Foroozanfard et al., 2011; Ganesh et al., 2009; Ghafarnegad et al., 2010; Ghanem et al., 2013; Hamed et al., 2010; Hassan et al., 2017; Hwu et al., 2005; Ibrahim et al., 2017; Lazovic et al., 1998; Legro et al., 2007; Legro et al., 2014; Liu et al., 2015; Machado et al., 2012; Malkawi and Qublan, 2002; Ng et al., 2001; Palomba et al., 2010; Palomba et al., 2005; Palomba et al., 2004; Rezk et al., 2018; Safdarian et al., 2012; Seyedoshohadaei et al., 2016; Sharma et al., 2010; Sohrabvand et al., 2006; Sturrock et al., 2002; Tasdemir et al., 2004; van Santbrink et al., 2005; Vandermolen et al., 2001; Vegetti et al., 1998; Weiss et al., 2017; Yadav et al., 2017; Yarali et al., 2002). The PRISMA flow diagram is presented in Figure 1 and the details of excluded studies are presented in Supplemental Table 2.

The included RCTs were published between 1998 and 2018, including two were conference abstracts (Lazovic et al., 1998; Sharma et al., 2010). Two RCTs were published in Persian (Ghafarnegad et al., 2010; Safdarian et al., 2012), one was in Portuguese (Machado et al., 2012) and the others were in English. Twelve RCTs included women with CC-failure PCOS (Abu Hashim et al., 2011a; Davar et al., 2011; De Leo et al., 1999; Ganesh et al., 2009; Ghafarnegad et al., 2010; Legro et al., 2007; Legro et al., 2014; Malkawi and Qublan, 2002; Palomba et al., 2005; Seyedoshohadaei et al., 2016; Sohrabvand et al., 2006; Weiss et al., 2017) while the others reported CC-resistant PCOS. The trial authors of two RCTs provided additional subset of data of women with CC-failure PCOS (Legro et al., 2007; Legro et al., 2014). Authors of another RCT provided additional outcome data (Bayram et al., 2004).

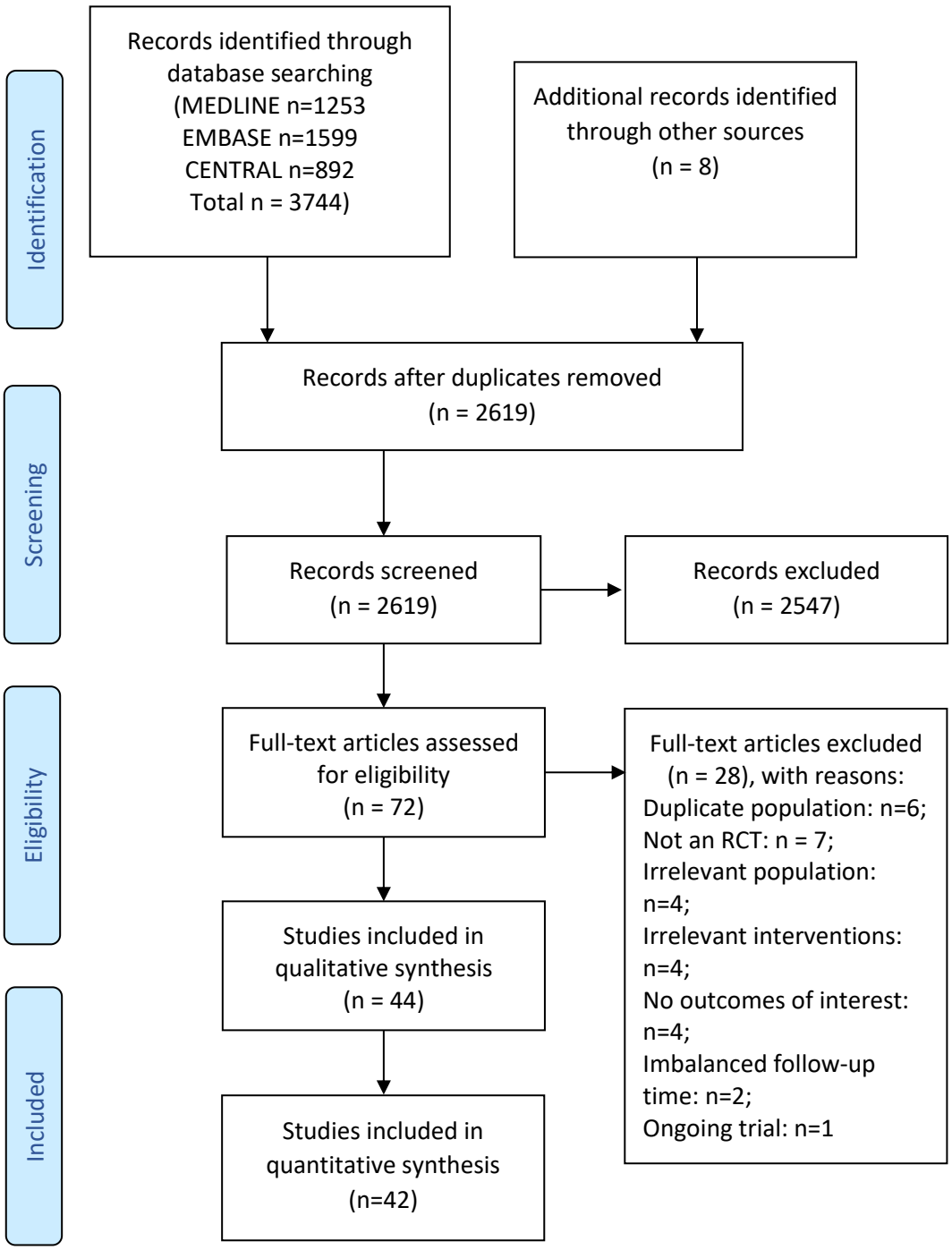


Figure 1 PRISMA flow diagram

The details of study characteristics are presented in Supplemental Table 3. Of 44 included RCTs, four were three-arm trials (Begum et al., 2013; Ganesh et al., 2009; Legro et al., 2007; Sharma et al., 2010) and the others were two-arm trials. These trials compared at least two of the following interventions to each other: CC, CC plus gonadotrophin, CC plus metformin, gonadotrophin, gonadotrophin plus metformin, letrozole, letrozole plus gonadotrophin, letrozole plus metformin, metformin and LOD.

Risk of bias of individual studies

The majority of included studies (64%, n=28) reported adequate methods of random sequence generation, while 48% (n=21) reported adequate methods of allocation concealment. Eleven studies were blinded trials and all studies were considered at low risk of performance bias due to the objective nature of the outcomes of interests. The majority of trials had low a risk of attrition bias (75%, n=33) while only approximately one third (36%, n=16) were considered at low risk of reporting bias. Risk of bias assessments of individual studies are presented in Supplemental Figure 1.

Network transitivity and consistency

A network meta-analysis relies on the transitivity assumption, which 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 this case, we consider the transitivity assumption valid as all included 10 interventions (CC, CC plus gonadotrophin, CC plus metformin, gonadotrophin, gonadotrophin plus metformin, letrozole, letrozole plus gonadotrophin, letrozole plus metformin, metformin and LOD) were valid treatment options for women treated unsuccessfully with CC. However, when including all 10 interventions in the network, we found evidence of global inconsistency (p= 0.0031). Local inconsistency test showed that such inconsistencies were due to the incoherence of the comparisons between metformin and

other interventions (CC versus metformin: p= 0.005; CC plus metformin versus metformin: p=0.013; LOD versus metformin: p<0.001).

Two RCTs comparing LOD versus metformin in CC-resistant women with PCOS showed conflicted results (Hamed et al., 2010; Palomba et al., 2004). In one RCT of 120 participants found that, compared to metformin, LOD resulted in less live births (OR 0.44, 95% CI 0.21-0.92) and no participants had a multiple pregnancy in either groups (Palomba et al., 2004). In the other RCT of 110 participants, live birth or multiple pregnancy was not reported and LOD showed higher clinical pregnancy rates than metformin (OR 2.47, 95% CI 1.05-5.81) (Hamed et al., 2010). In a subset (n=310) of CC-failure women with PCOS in a three-arm RCT comparing CC plus metformin, CC alone and metformin alone, both CC plus metformin and CC alone showed higher odds of live birth compared to metformin alone (OR 4.71, 95%CI 1.82-12.20; OR 4.32, 95% CI 1.67-11.16) (Legro et al., 2007).

Given the observed inconsistency in the network, we excluded metformin arm from the network, i.e. we excluded two RCTs comparing LOD versus metformin and the metformin arm in a three-arm RCT. The remaining 42 RCTs involving 6925 women were included in network meta-analyses and no evidence of inconsistency was observed.

Live birth / ongoing pregnancy

Twenty-one RCTs reported live birth/ongoing pregnancy in 3,072 women and these RCTs compared the following eight interventions to each other: CC, letrozole, gonadotrophins, CC plus metformin, letrozole plus metformin, gonadotrophins plus metformin, CC plus gonadotrophins and laparoscopic ovarian drilling (Figure 2).

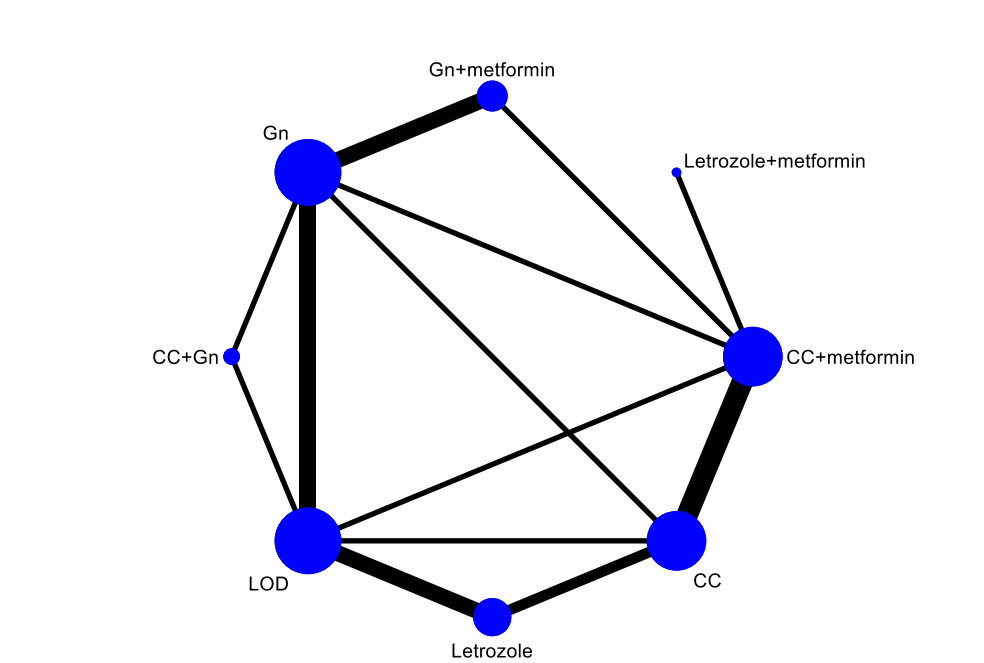


Figure 2 Network plot for live birth/ongoing pregnancy

Network meta-analysis (Figure 3) showed that compared to CC, letrozole alone and gonadotrophins alone resulted in higher odds of live birth/ongoing pregnancy (OR 1.67, 95% CI 1.16-2.40, moderate-certainty evidence; and OR 1.63, 95% CI 1.24-2.14, high-certainty evidence, respectively), while the addition of metformin to these two interventions improved live birth/ongoing pregnancy further (OR 4.89, 95% CI 1.10-21.81, low-certainty evidence; and OR 4.47, 95% CI 2.37-8.41, moderate-certainty evidence, respectively). The evidence of difference between CC plus metformin, CC plus gonadotrophins or LOD versus CC was insufficient (OR 1.09, 95% CI 0.67-1.77; OR 1.14, 95% CI 0.81-1.61; OR 1.81, 95% CI 0.97-3.40; all low-certainty evidence). Results of pairwise meta-analyses were consistent with those of network meta-analysis in terms of letrozole or gonadotrophins versus CC (Supplemental Table 4).

SUCRA values for gonadotrophins plus metformin, letrozole plus metformin, CC plus gonadotrophins, letrozole, gonadotrophins, LOD, CC plus metformin and CC were 92.2%, 88.2%, 58.7%, 57.1%, 54.4%, 21.9%, 18.3% and 9.2%, respectively (Supplemental Figure 2). The results in pairwise meta-analyses were consistent with those in network meta-analysis.

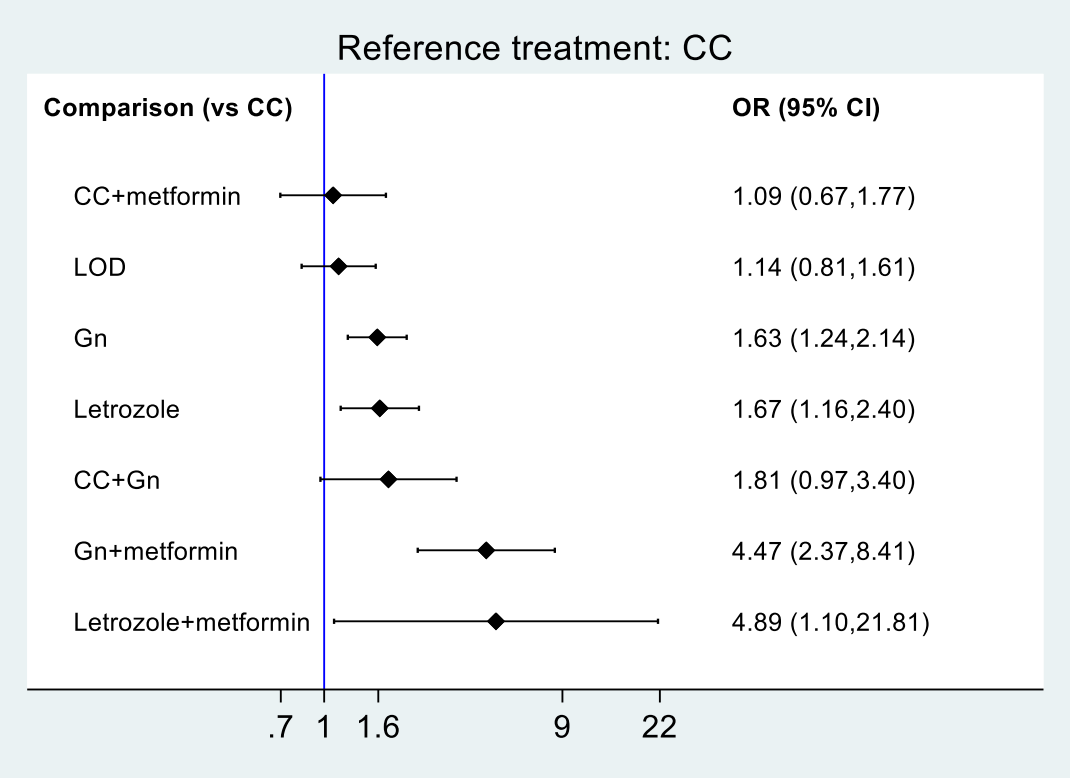


Figure 3 Network meta-analysis for live birth/ongoing pregnancy.
Diamonds and lines represent odds ratios (ORs) and relevant 95% confidence intervals (CIs), respectively. OR<1 favours CC; OR>1 favours other interventions.

Subgroup analysis and sensitivity analysis

Subgroup analysis on RCTs of CC-failure PCOS (8 RCTs, 1741 women) showed similar results on live birth/ongoing pregnancy (Supplemental Table 5). In women with CC-resistant PCOS (13 RCTs, 1331 women), apart from letrozole plus metformin, all alternative interventions, including LOD seemed to be superior to CC (Supplemental Table 5). The effect sizes were larger compared to those in the main analysis and the confidence intervals were overall very wide due to the small sample sizes of most studies.

Sensitivity analysis on RCTs with low risk of bias at both random sequence generation and allocation concealment (12 RCTs, 2350 women) showed similar results on live birth/ongoing pregnancy (Supplemental Table 6).

Multiple pregnancy

Twenty RCTs with 3944 women were included in the network meta-analysis comparing the following eight interventions to each other: CC, CC plus gonadotrophins, CC plus

metformin, gonadotrophins, gonadotrophins plus metformin, LOD, letrozole and letrozole plus gonadotrophins (Supplemental Figure 3). Network meta-analysis (Supplemental Table 7) showed that compared to CC, LOD resulted in lower odds of multiple pregnancy (OR 0.13, 95% CI 0.03-0.51, moderate-certainty evidence). There was insufficient evidence of a difference between the other interventions and CC. SUCRA values for gonadotrophins, CC, letrozole, gonadotrophins plus metformin, CC plus gonadotrophins, CC plus metformin, letrozole plus gonadotrophin, gonadotrophins, and LOD were 83.4%, 79.3%, 57.6%, 55.6%, 53.8%, 42.9%, 19.1% and 8.3%, respectively.

Secondary outcomes

Forty-one RCTs with 6836 women reported clinical pregnancy (Supplemental Figure 4). Network meta-analysis (Supplemental Table 7) showed that compared to CC alone, gonadotrophins plus metformin, letrozole, letrozole plus metformin, gonadotrophins, letrozole plus gonadotrophins, CC plus gonadotrophins and LOD resulted in higher odds of clinical pregnancy. The evidence of a difference between CC plus metformin versus CC was insufficient (OR 1.33, 95% CI 0.90-1.95). SUCRA values for gonadotrophins plus metformin, letrozole, gonadotrophins, letrozole plus metformin, letrozole plus gonadotrophins, CC plus gonadotrophins, LOD, CC plus metformin and CC were 99.7%, 71%, 66.9%, 62.9%, 61.9%, 38.3%, 28.8%, 18.7% and 1.8%, respectively.

Twenty-three RCTs including 3821 women reported ovulation (Supplemental Figure 5). Network meta-analysis (Supplemental Table 7) showed that compared to CC alone, gonadotrophins plus metformin, gonadotrophins, letrozole, letrozole plus metformin, LOD, CC plus metformin resulted in higher odds of ovulation. The evidence of a difference between CC plus gonadotrophins versus CC was insufficient (OR 2.04, 95% CI 0.71-5.89). SUCRA values for gonadotrophins plus metformin, letrozole, letrozole plus metformin, gonadotrophins, LOD, CC plus metformin, CC plus gonadotrophins and CC were 97%, 88.3%, 69.6%, 40%, 38.5%, 36.6%, 27.1% and 2.9%, respectively.

Thirty-three RCTs including 6362 women reported miscarriage (Supplemental Figure 6). Compared to CC, gonadotrophins resulted in more miscarriages (OR 2.22, 95% CI 1.28-3.84). There was insufficient evidence of a difference between other interventions and CC on miscarriage (Supplemental Table 7).

Meta-analysis on OHSS was not performed due to the small number of events, especially in studies comparing non-gonadotrophin interventions. Twelve RCTs reported no OHSS (Abu Hashim et al., 2011a; Abu Hashim et al., 2010a; Abu Hashim et al., 2010b; Abu Hashim et al., 2011b; Bayram et al., 2004; Foroozanfard et al., 2011; Ghanem et al., 2013; Hamed et al., 2010; Hassan et al., 2017; Liu et al., 2015; Vegetti et al., 1998; Weiss et al., 2017). One RCT reported two cases of mild OHSS (2/78) in the gonadotrophin group and no case of OHSS (0/75) in the CC plus metformin group (Abu Hashim et al., 2011c). One RCT reported two cases of OHSS (2/12) in the CC group and no case of OHSS (0/16) in the CC plus metformin group (Malkawi and Qublan, 2002). Another RCT reported one case of OHSS in the CC plus gonadotrophins (1/35) and no case of OHSS in the gonadotrophins group (0/35) (Palomba et al., 2005).

DISCUSSION

Summary of key findings

Our systematic review showed in women with PCOS treated unsuccessfully with CC, letrozole or gonadotrophins alone, or in combination with metformin, are the top four ranking interventions for live birth, clinical pregnancy and ovulation. These effects were observed in both CC-resistant and CC-failure women with PCOS and therefore these interventions should be considered as the second-line ovulation induction options. The conclusion on the adding value of metformin should be interpreted with caution due to the small number of included studies and concerns on the risk of bias.

LOD resulted in more live birth in women with CC-resistant PCOS but the evidence of a difference was insufficient in CC-failure PCOS. LOD resulted in the lowest multiple pregnancy rate while the evidence on the other interventions versus CC on multiple pregnancy was insufficient.

Strengths and limitations

The strengths of this systematic review included extensive search strategies without language limitations, the use of both direct and indirect evidence, presentation of the hierarchy of rankings of effectiveness to guide clinical decision. In addition, we reported the overall certainty of evidence for easier interpretation.

Several limitations should be considered during data interpretation. The primary effectiveness (live birth / ongoing pregnancy) and safety (multiple pregnancy) outcomes were only reported in 50% of the included trials. The quality of studies not reporting these primary outcomes had relatively lower quality compared to those reporting these outcomes. Among the RCTs reporting the primary outcome, about 50% did not report the details of random sequence generation or allocation concealment, which restricts the interpretation of the validity of these results. In addition, we did not differentiate the dosage of each ovulation induction medication but focused on the comparisons of different medications. Moreover,

we only focused on the effect of LOD itself on fertility outcomes and did not consider any subsequent ovulation induction with pharmacological therapies after LOD. Therefore, a strategy involving LOD plus subsequent ovulation induction was beyond our research question. Finally, the variations of the operative techniques during LOD were not considered in the analyses, which may partly explain the heterogeneous effects of LOD in different RCTs.

Clinical implications

Letrozole, gonadotrophins, letrozole plus metformin and gonadotrophins plus metformin can be considered as the second-line ovulation induction strategies in women with PCOS who are treated with CC unsuccessfully. As letrozole is off-label use for ovulation induction, it can only be used in settings where off-label use is allowed, and women are fully informed. Costs, risks, complexity of monitoring and administration, and issues of off-label use should be considered in the context of shared-decision making.

Based on our network meta-analysis, LOD itself is not considered as a top-ranking ovulation induction method as a second-line ovulation induction for PCOS. For women with CC-resistant PCOS, LOD increases live birth rate but the effect may not exist for women with CC-failure PCOS.

Research implications

The adding value of metformin to existing ovulation induction medications such as letrozole and gonadotrophins should be further evaluated in RCTs, given the limited number of studies comparing these interventions with concerns on risk of bias. New trials are also encouraged to incorporate treatment selection markers such as biomarkers of insulin resistance in their design to guide personalised treatment decision (Wang et al., 2019).

Letrozole has been recommended as the first-line treatment in the latest international evidence-based guideline (Teede et al., 2018), but the treatment choices for women who are

treated unsuccessfully with letrozole are unclear. The top-ranking interventions in this network meta-analysis can be used as an evidence-base to design new trials on women with letrozole-resistant or letrozole-failure PCOS.

CONCLUSIONS

Our systematic review and network meta-analysis showed in women with PCOS treated unsuccessfully with CC, letrozole or gonadotrophins alone, or in combination with metformin, are the top-ranking interventions for live birth, clinical pregnancy and ovulation. These effects were observed in both CC-resistant and CC-failure women with PCOS and therefore these interventions should be considered as the second-line ovulation induction options.

REFERENCES

Abd Elgafor I. Efficacy of combined metformin-letrozole in comparison with bilateral ovarian drilling in clomiphene-resistant infertile women with polycystic ovarian syndrome. *Archives of Gynecology & Obstetrics* 2013;288: 119-123.

Abdellah MS. Reproductive outcome after letrozole versus laparoscopic ovarian drilling for clomiphene-resistant polycystic ovary syndrome. *International Journal of Gynaecology & Obstetrics* 2011;113: 218-221.

Abu Hashim H, Foda O, Ghayaty E. Combined metformin-clomiphene in clomiphene-resistant polycystic ovary syndrome: a systematic review and meta-analysis of randomized controlled trials. *Acta Obstet Gynecol Scand* 2015;94: 921-930.

Abu Hashim H, Foda O, Ghayaty E, Elawa A. Laparoscopic ovarian diathermy after clomiphene failure in polycystic ovary syndrome: is it worthwhile? A randomized controlled trial. 2011a, pp. 1303-1309.

Abu Hashim H, Mashaly AM, Badawy A. Letrozole versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial. *Archives of Gynecology & Obstetrics* 2010a;282: 567-571.

Abu Hashim H, Shokeir T, Badawy A. Letrozole versus combined metformin and clomiphene citrate for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: a randomized controlled trial. *Fertility & Sterility* 2010b;94: 1405-1409.

Abu Hashim H, Wafa A, El Rakhawy M. Combined metformin and clomiphene citrate versus highly purified FSH for ovulation induction in clomiphene-resistant PCOS women: a randomised controlled trial. *Gynecological Endocrinology* 2011b;27: 190-196.

Abu Hashim HA, Lakany NE, Sherief L. Combined metformin and clomiphene citrate versus laparoscopic ovarian diathermy for ovulation induction in clomiphene-resistant women with polycystic ovary syndrome: A randomized controlled trial. *Journal of Obstetrics and Gynaecology Research* 2011c;37: 169-177.

Balen AH, Morley LC, Misso M, Franks S, Legro RS, Wijeyaratne CN, Stener-Victorin E, Fauser BC, Norman RJ, Teede H. The management of anovulatory infertility in women with polycystic ovary syndrome: an analysis of the evidence to support the development of global WHO guidance. *Hum Reprod Update* 2016;22: 687-708.

Bayram N, Wely M, Kaaijk E, Bossuyt P, Veen F. Using an electrocautery strategy or recombinant follicle stimulating hormone to induce ovulation in polycystic ovary syndrome: randomised controlled trial. 2004, pp. 192.

Begum MR, Akhter S, Ehsan M, Begum MS, Khan F. Pretreatment and co-administration of oral anti-diabetic agent with clomiphene citrate or rFSH for ovulation induction in clomiphene-citrate-resistant polycystic ovary syndrome. *Journal of Obstetrics & Gynaecology Research* 2013;39: 966-973.

Begum MR, Ferdous J, Begum A, Quadir E. Comparison of efficacy of aromatase inhibitor and clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. *Fertility & Sterility* 2009;92: 853-857.

Bordewijk EM, Nahuis M, Costello MF, Van der Veen F, Tso LO, Mol BW, van Wely M. Metformin during ovulation induction with gonadotrophins followed by timed intercourse or intrauterine insemination for subfertility associated with polycystic ovary syndrome. *Cochrane Database Syst Rev* 2017;1: CD009090.

Chaimani A, Salanti G. Visualizing Assumptions and Results in Network Meta-analysis: The Network Graphs Package. *The Stata Journal* 2015;15: 905-950.

Davar R, Javedani M, Fallahzadeh MH. Metformin-letrozole in comparison with Metformin-clomiphene citrate in clomiphene-resistance PCOS patients undergoing IUI. *Iranian Journal of Reproductive Medicine* 2011;9: 31-36.

De Leo V, Marca AL, Ditto A, Morgante G, Cianci A. Effects of metformin on gonadotropin-induced ovulation in women with polycystic ovary syndrome. *Fertility and Sterility* 1999;72: 282-285.

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

Farquhar C, Brown J, Marjoribanks J. Laparoscopic drilling by diathermy or laser for ovulation induction in anovulatory polycystic ovary syndrome. *Cochrane Database Syst Rev* 2012: CD001122.

Farquhar CM, Williamson K, Gudex G, Johnson NP, Garland J, Sadler L. A randomized controlled trial of laparoscopic ovarian diathermy versus gonadotropin therapy for women with clomiphene citrate-resistant polycystic ovary syndrome. *Fertility & Sterility* 2002;78: 404-411.

Foroozanfard F, Mahdian M, Mousavi G, Pejmanmanesh M, Soleimani A. Use of letrozole versus clomiphsene citrate combined with gonadotropins in patients with clomiphene resistant polycystic ovarian syndrome: A comparative study. *Pakistan Journal of Medical Sciences* 2011;27: 967-970.

Franik S, Eltrop SM, Kremer JA, Kiesel L, Farquhar C. Aromatase inhibitors (letrozole) for subfertile women with polycystic ovary syndrome. *Cochrane Database Syst Rev* 2018;5: CD010287.

Ganesh A, Goswami SK, Chattopadhyay R, Chaudhury K, Chakravarty B. Comparison of letrozole with continuous gonadotropins and clomiphene-gonadotropin combination for ovulation induction in 1387 PCOS women after clomiphene citrate failure: a randomized prospective clinical trial. *Journal of Assisted Reproduction & Genetics* 2009;26: 19-24.

Ghafarnegad M, Arjmand N, Khazaeipour Z. Pregnancy rate of gonadotrophin therapy and laparoscopic ovarian electrocautery in polycystic ovary syndrome resistant to clomiphene citrate: A comparative study. *Tehran University Medical Journal* 2010;67: 712-717.

Ghanem ME, Elboghdady LA, Hassan M, Helal AS, Gibreel A, Houssen M, Shaker ME, Bahlol I, Mesbah Y. Clomiphene Citrate co-treatment with low dose urinary FSH versus urinary FSH for clomiphene resistant PCOS: Randomized controlled trial. *Journal of Assisted Reproduction and Genetics* 2013;30: 1477-1485.

Hamed HO, Hasan AF, Ahmed OG, Ahmed MA. Metformin versus laparoscopic ovarian drilling in clomiphene- and insulin-resistant women with polycystic ovary syndrome. *International Journal of Gynaecology & Obstetrics* 2010;108: 143-147.

Hassan A, Shehata N, Wahba A. Cost effectiveness of letrozole and purified urinary FSH in treating women with clomiphene citrate-resistant polycystic ovarian syndrome: a randomized controlled trial. *Human Fertility* 2017;20: 37-42.

Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011: Available from www.handbook.cochrane.org.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162: 777-784.

Hwu YM, Lin SY, Huang WY, Lin MH, Lee RKK. Ultra-short metformin pretreatment for clomiphene citrate-resistant polycystic ovary syndrome. *International Journal of Gynecology and Obstetrics* 2005;90: 39-43.

Ibrahim MH, Tawfic M, Hassan MM, Sedky OH. Letrozole versus laparoscopic ovarian drilling in infertile women with PCOS resistant to clomiphene citrate. *Middle East Fertility Society Journal* 2017;22: 251-254.

Lazovic G, Milacic D, Terzic M, Spremovic S, Mitijasevic S. Medicaments or surgical therapy of PCOS (Abstract only). *Fertility and Sterility 1998;70(3):472* 1998;70: S472.

Legro RS, Barnhart HX, Schlaff WD, Carr BR, Diamond MP, Carson SA, Steinkampf MP, Coutifaris C, McGovern PG, Cataldo NA *et al.* Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. *N Engl J Med* 2007;356: 551-566.

Legro RS, Brzyski RG, Diamond MP, Coutifaris C, Schlaff WD, Casson P, Christman GM, Huang H, Yan Q, Alvero R *et al.* Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. *N Engl J Med* 2014;371: 119-129.

Liu W, Dong S, Li Y, Shi L, Zhou W, Liu Y, Liu J, Ji Y. Randomized controlled trial comparing letrozole with laparoscopic ovarian drilling in women with clomiphene citrate-resistant polycystic ovary syndrome. *Experimental and Therapeutic Medicine* 2015;10: 1297-1302.

Machado RC, Machado NA, Geber S. Evaluation of the use of metformin in ovulation result of patients with the polycystic ovary syndrome resistant to isolated use of clomiphene citrate. *Jornal Brasileiro de Reproducao Assistida* 2012;16: 27-31.

Malkawi HY, Qublan HS. The effect of metform plus clomiphene citrate on ovulation and pregnancy rates in clomiphene-resistant women with polycystic ovary syndrome. *Saudi Medical Journal* 2002;23: 663-666.

Morley LC, Tang T, Yasmin E, Norman RJ, Balen AH. Insulin-sensitising drugs (metformin, rosiglitazone, pioglitazone, D-chiro-inositol) for women with polycystic ovary syndrome, oligo amenorrhoea and subfertility. *Cochrane Database Syst Rev* 2017;11: CD003053.

Ng EHY, Wat NMS, Ho PC. Effects of metformin on ovulation rate, hormonal and metabolic profiles in women with clomiphene-resistant polycystic ovaries: A randomized, double-blinded placebo-controlled trial. *Human Reproduction* 2001;16: 1625-1631.

Nikolakopoulou A, Higgins JP, Papakonstantinou T, Chaimani A, Del Giovane C, Egger M, Salanti G. Assessing Confidence in the Results of Network Meta-Analysis (Cinema). *bioRxiv* 2019: 597047.

Palomba S, Falbo A, Battista L, Russo T, Venturella R, Tolino A, Orio F, Zullo F. Laparoscopic ovarian diathermy vs clomiphene citrate plus metformin as second-line strategy for infertile anovulatory patients with polycystic ovary syndrome: a randomized controlled trial. *American Journal of Obstetrics & Gynecology* 2010;202: 577.e571-578.

Palomba S, Falbo A, Orio Jr F, Manguso F, Russo T, Tolino A, Colao A, Dale B, Zullo F. A randomized controlled trial evaluating metformin pre-treatment and co-administration in non-obese insulin-resistant women with polycystic ovary syndrome treated with controlled

ovarian stimulation plus timed intercourse or intrauterine insemination. *Human Reproduction* 2005;20: 2879-2886.

Palomba S, Orio Jr F, Nardo LG, Falbo A, Russo T, Corea D, Doldo P, Lombardi G, Tolino A, Colao A *et al.* Metformin administration versus laparoscopic ovarian diathermy in clomiphene citrate-resistant women with polycystic ovary syndrome: A prospective parallel randomized double-blind placebo-controlled trial. *Journal of Clinical Endocrinology and Metabolism* 2004;89: 4801-4809.

Rezk M, Shaheen AE, Saif El-Nasr I. Clomiphene citrate combined with metformin versus letrozole for induction of ovulation in clomiphene-resistant polycystic ovary syndrome: a randomized clinical trial. *Gynecological Endocrinology* 2018;34: 298-300.

Riley RD, Jackson D, Salanti G, Burke DL, Price M, Kirkham J, White IR. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017;358: j3932.

Safdarian L, Dibazar NS, Ahmadzadeh A, Yekta BG. Comparing two ovulation induction methods by brachial artery ultrasonography in infertile women with polycystic ovary syndrome. *Tehran University Medical Journal* 2012;70: 242-249.

Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64: 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.

Seyedoshohadaei F, Tangestani L, Zandvakili F, Rashadmanesh N. Comparison of the effect of clomiphene- estradiol valerate vs letrozole on endometrial thickness, abortion and pregnancy rate in infertile women with polycystic ovarian syndrome. *Journal of Clinical and Diagnostic Research* 2016;10: QC10-QC13.

Sharma S, Goswami SK, Rajani S, Mukherjee S, Modi R, Sarkar A, Chakravarty BN. Effcacy of letrozole-gonadotropin, clomiphene-gonadotropin and only gonadotropin in

clomiphene and letrozole resistant PCOS women undergoing IUI. *Human Reproduction* 2010;25: i46.

Sohrabvand F, Ansari S, Bagheri M. Efficacy of combined metformin-letrozole in comparison with metformin-clomiphene citrate in clomiphene-resistant infertile women with polycystic ovarian disease. *Human Reproduction* 2006;21: 1432-1435.

Sturrock NDC, Lannon B, Fay TN. Metformin does not enhance ovulation induction in clomiphene resistant polycystic ovary syndrome in clinical practice. *British Journal of Clinical Pharmacology* 2002;53: 469-473.

Tasdemir S, Ficicioglu C, Yalti S, Gurbuz B, Basaran T, Yildirim G. The effect of metformin treatment to ovarian response in cases with PCOS. *Archives of Gynecology & Obstetrics* 2004;269: 121-124.

Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ, International PN. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Fertil Steril* 2018;110: 364-379.

van Santbrink EJ, Hohmann FP, Eijkemans MJ, Laven JS, Fauser BC. Does metformin modify ovarian responsiveness during exogenous FSH ovulation induction in normogonadotrophic anovulation? A placebo-controlled double-blind assessment. *Eur J Endocrinol* 2005;152: 611-617.

Vandermolen DT, Ratts VS, Evans WS, Stovall DW, Kauma SW, Nestler JE. Metformin increases the ovulatory rate and pregnancy rate from clomiphene citrate in patients with polycystic ovary syndrome who are resistant to clomiphene citrate alone. *Fertility & Sterility* 2001;75: 310-315.

Vegetti W, Ragni G, Baroni E, Testa G, Marsico S. Laparoscopic ovarian drilling versus low-dose pure FSH in anovulatory clomiphene-resistant patients with polycystic ovarian syndrome: randomized prospective study. 1998, pp. 120.

Wang R, Li W, Bordewijk EM, Legro RS, Zhang H, Wu X, Gao J, Morin-Papunen L, Homburg R, König TE *et al.* First-line ovulation induction for polycystic ovary syndrome: an individual participant data meta-analysis. *Hum Reprod Update* 2019;in press.

Weiss NS, Kostova E, Nahuis M, Mol BWJ, van der Veen F, van Wely M. Gonadotrophins for ovulation induction in women with polycystic ovary syndrome. *Cochrane Database Syst Rev* 2019;1: CD010290.

Weiss NS, Nahuis MJ, Bordewijk E, Oosterhuis JE, Smeenk JMJ, Hoek A, Broekmans FJM, Fleischer K, de Bruin JP, Kaaijk EM *et al.* Gonadotrophins versus clomifene citrate with or without intrauterine insemination in women with normogonadotropic anovulation and clomifene failure (M-OVIN): A randomised, two-by-two factorial trial. *The Lancet* 2017.

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

Yadav P, Singh S, Singh R, Jain M, Awasthi S, Raj P. To study the effect on fertility outcome by gonadotropins vs laparoscopic ovarian drilling in clomiphene-resistant cases of polycystic ovarian syndrome. *Journal of SAFOG* 2017;9: 336-340.

Yarali H, Yildiz BO, Demirol A, Zeyneloğlu HB, Yiğit N, Bükülmez O, Koray Z. Co-administration of metformin during rFSH treatment in patients with clomiphene citrate-resistant polycystic ovarian syndrome: A prospective randomized trial. *Human Reproduction* 2002;17: 289-294.

Yu Y, Fang L, Zhang R, He J, Xiong Y, Guo X, Du Q, Huang Y, Sun Y. Comparative effectiveness of 9 ovulation-induction therapies in patients with clomiphene citrate-resistant polycystic ovary syndrome: a network meta-analysis. *Sci Rep* 2017;7: 3812.

Chapter 5

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

This is the accepted version of a Cochrane Review 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 of Systematic Reviews* 2019,9:CD012692. doi: 10.1002/14651858.CD012692.pub2, which has been published in final form at: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD012692.pub2/full

Statement of Authorship

Title of Paper	Interventions for unexplained infertility: a systematic review and network meta-analysis
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	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 of Systematic Reviews 2019, 9:CD012692. doi: 10.1002/14651858.CD012692.pub2.

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	RW contributed to the study design and the protocol development, collected the data, analysed the data, interpreted the data, wrote the first draft and approved the final version.		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	06/09/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

iv. the candidate's stated contribution to the publication is accurate (as detailed above);

v. permission is granted for the candidate in include the publication in the thesis; and

vi. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Nora A Danhof		
Contribution to the Paper	NAD collected the data, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	06/09/2019

Name of Co-Author	Raissa I Tjon-Kon-Fat
Contribution to the Paper	RIT contributed to the study design and the protocol development collected the data, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.

Signature		Date	06/09/2019

Name of Co-Author	Marinus JC Eijkemans		
Contribution to the Paper	MJCE contributed to the study design and the protocol development, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	9/9/2019

Name of Co-Author	Patrick MM Bossuyt		
Contribution to the Paper	PMMB contributed to the study design and the protocol development, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	9/9/2019

Name of Co-Author	Monique H Mochtar		
Contribution to the Paper	NAD interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	10/9/2019

Name of Co-Author	Fulco van der Veen		
Contribution to the Paper	FvdV contributed to the study design and the protocol development, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	9/9/2019

Name of Co-Author	Siladitya Bhattacharya		
Contribution to the Paper	SB contributed to the study design and the protocol development, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	9/9/2019

Name of Co-Author	Ben Willem J Mol		
Contribution to the Paper	BWJM contributed to the study design and the protocol development, interpreted the data, revised the manuscript critically for important intellectual content and approved the final version.		
Signature		Date	6/9/2019

Name of Co-Author	Madelon van Wely		
Contribution to the Paper	MvW contributed to the study design and the protocol development, collected the data, analysed the data, interpreted the data, wrote the first draft, approved the final version and acted as the corresponding author.		
Signature		Date	6/9/2019

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

Review question

Researchers in The Cochrane Collaboration 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 * 10^6 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.
  - www.clinicaltrials.gov (a service of the US National Institutes of Health).
  - 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(ORAvsB) = ln(ORAvsC) − ln(ORBvsC). 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., 1991b; 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., 2014a; 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., 2014a; 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., 1991b; 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., 2014a; 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., 1991a).

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., 1991b); 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., 2014a), 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., 1991b; 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., 2014a; 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., 1991b) 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., 1991b; 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., 2014a; 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).

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., 2014b; 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).

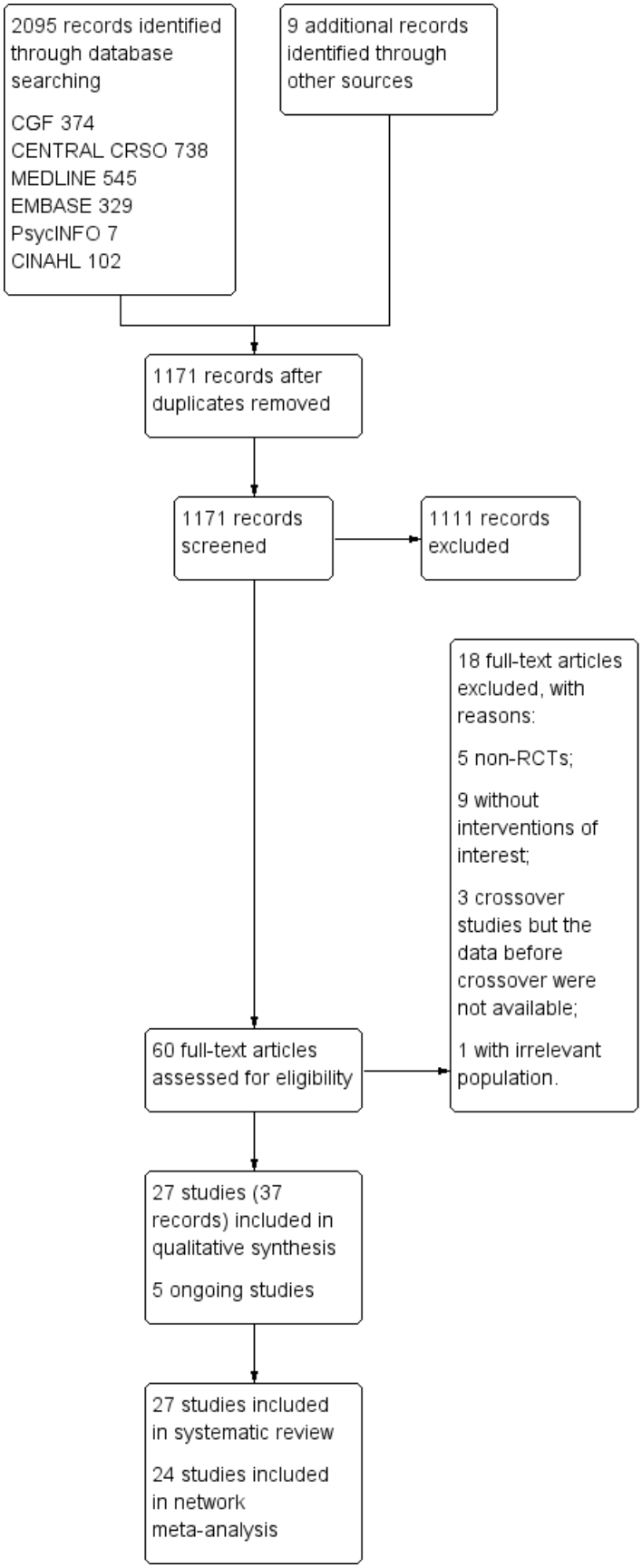


Figure 1 Study flow diagram.

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., 2014a). 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.

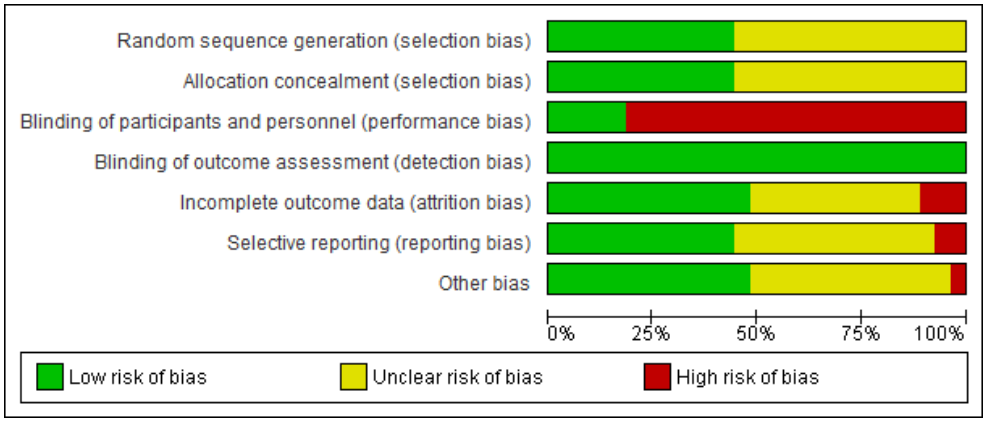


Figure 2 Risk of bias graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

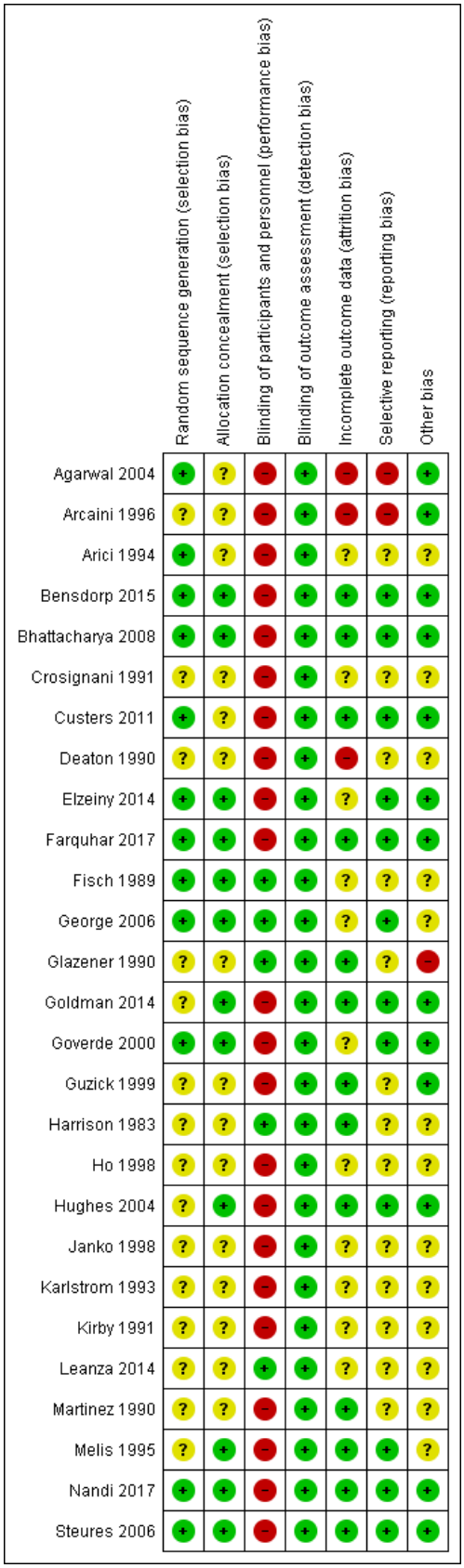


Figure 3 Risk of bias summary: review authors' judgements about each risk of bias item for each included study.

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.

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., 1991b; 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., 1991b; George et al., 2006; Ho et al., 1998; Janko et al., 1998; Karlstrom et al., 1993; Kirby et al., 1991; Leanza et al., 2014a). 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.

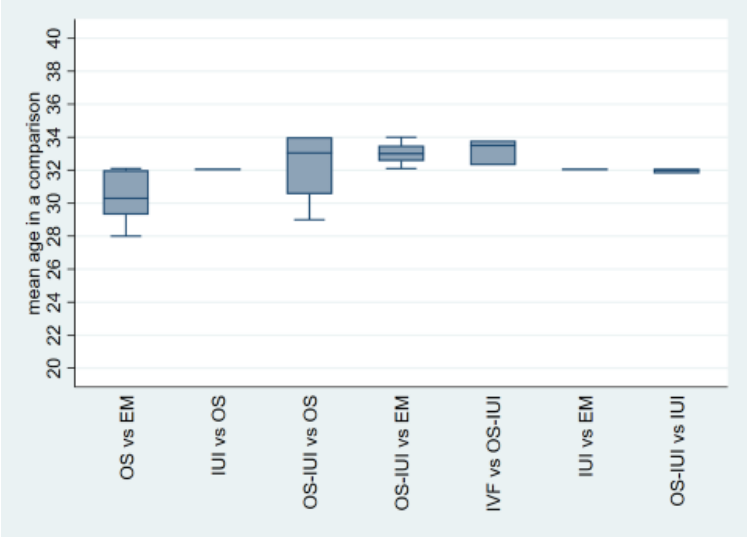


Figure 4 Box plot for the distribution of means of age in different studies across different comparisons.

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).

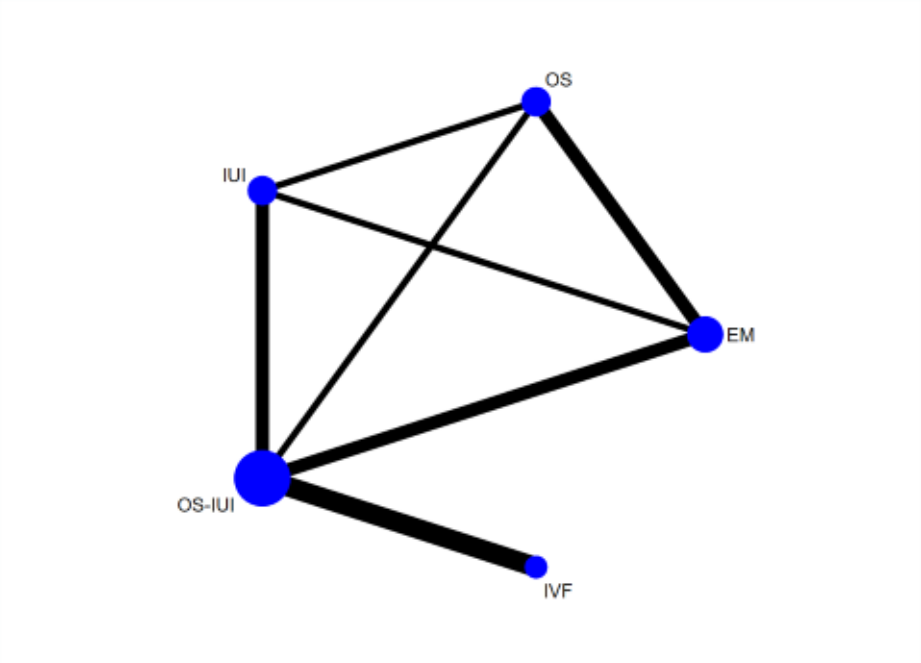


Figure 5 Network plot for live birth.
Each node represents an intervention, and the size of each node is proportional to the number of trials reporting such intervention. The widths of the lines are proportional to the numbers of trials comparing each pair of interventions.

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.

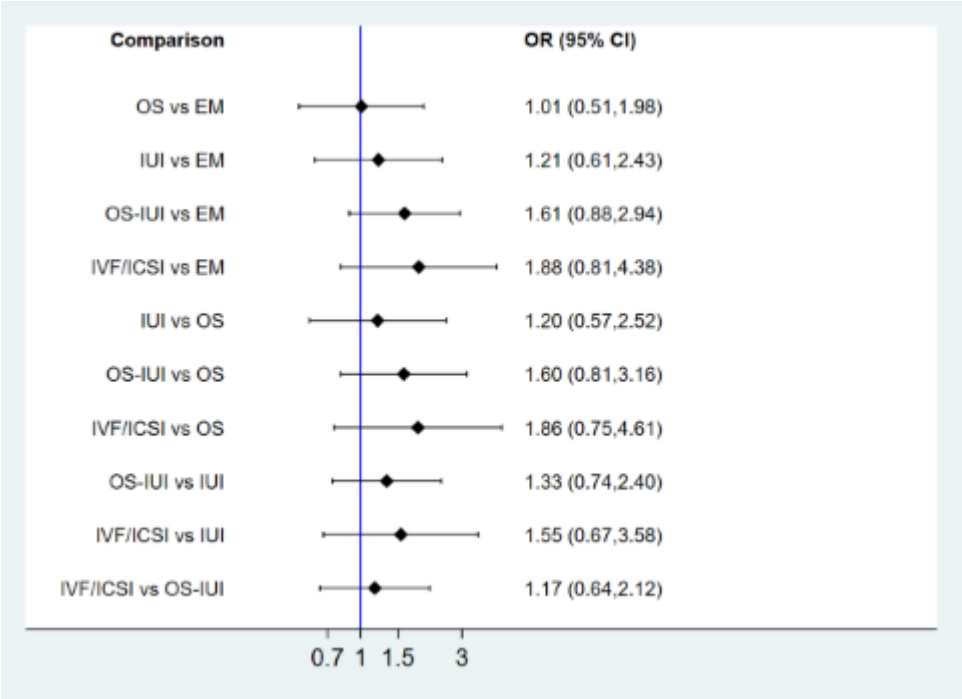


Figure 6 Network meta-analysis for live birth.
Each diamond represents the estimate summary odds ratio of each comparison; each horizontal line represents the confidence interval of each comparison; blue vertical line represents line of no effect (odds ratio = 1). Odds ratio greater than 1 favours the 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 (Appendix 7.1). 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.

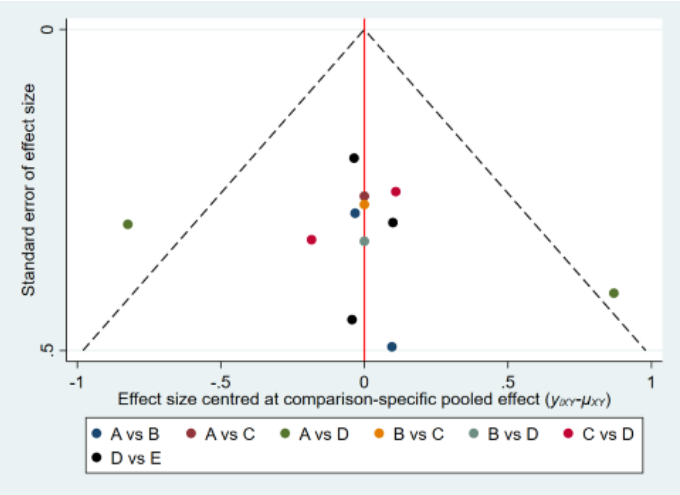


Figure 7 Comparison-adjusted funnel plot for live birth.
(A: expectant management; B: OS; C: IUI; D: OS-IUI; E: IVF/ICSI.)

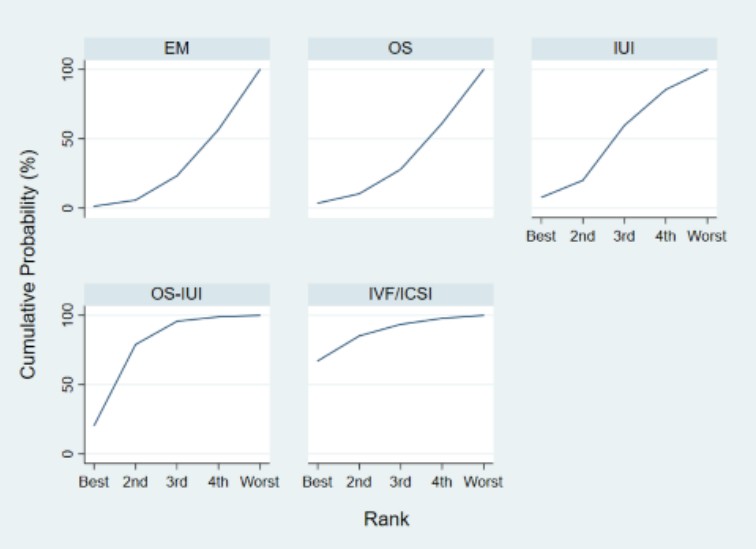


Figure 8 Cumulative rankograms of interventions for live birth.
Each cumulative rankogram illustrates the cumulative probability of each ranking (from the best to the worst rank) for each intervention in terms of live birth.

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.

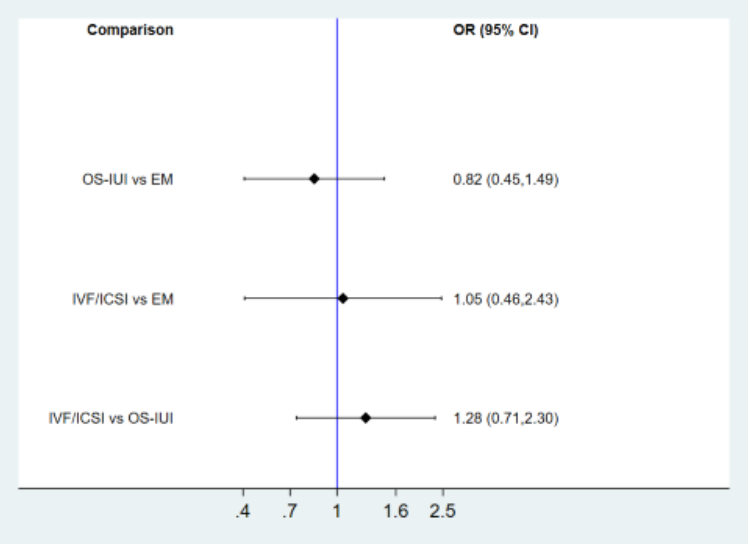


Figure 9 Subgroup analysis for live birth - RCTs with a median duration of infertility ≤ 2 years.

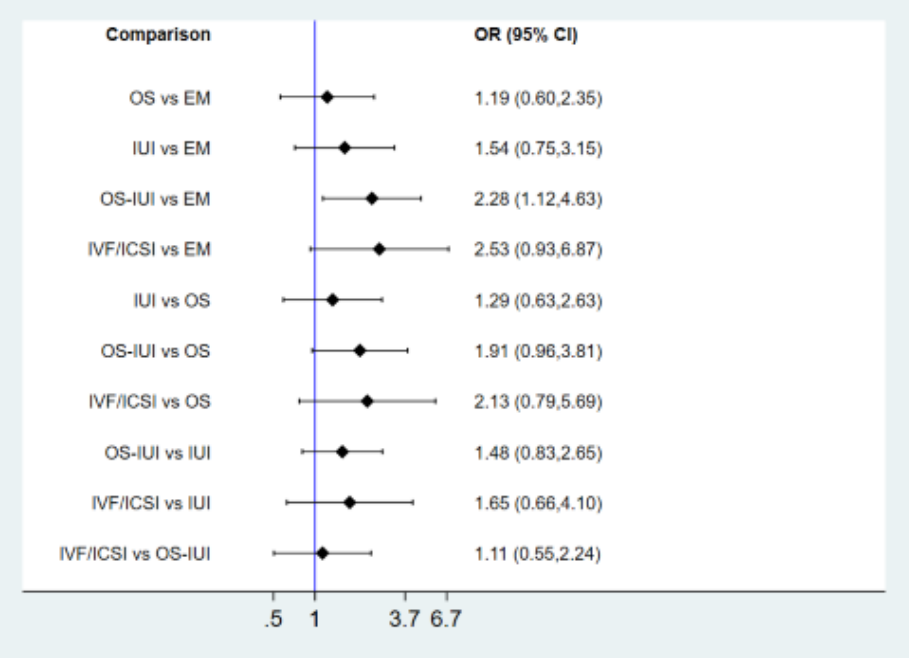


Figure 10 Subgroup analysis for live birth - RCTs with a median duration of infertility >2 years.

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.

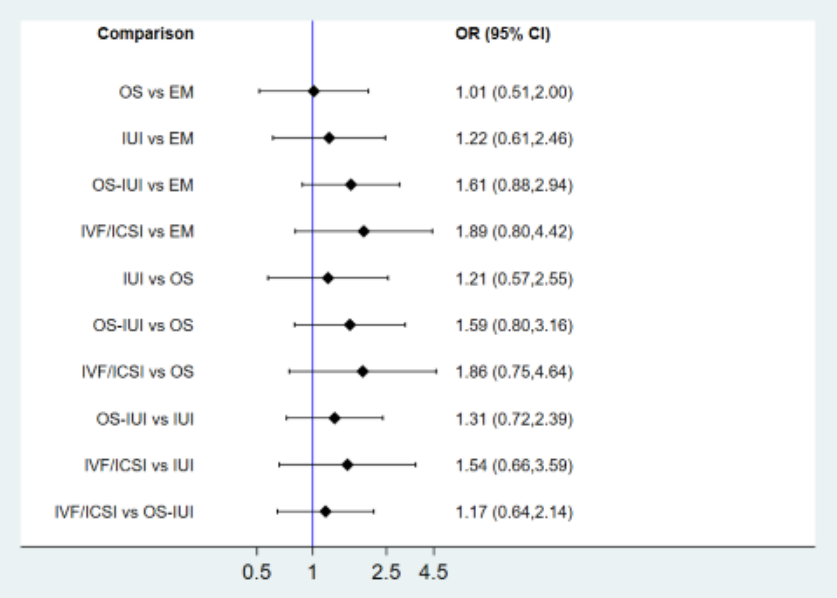


Figure 11 Sensitivity analysis for live birth by exclusion of participants with missing outcome data.

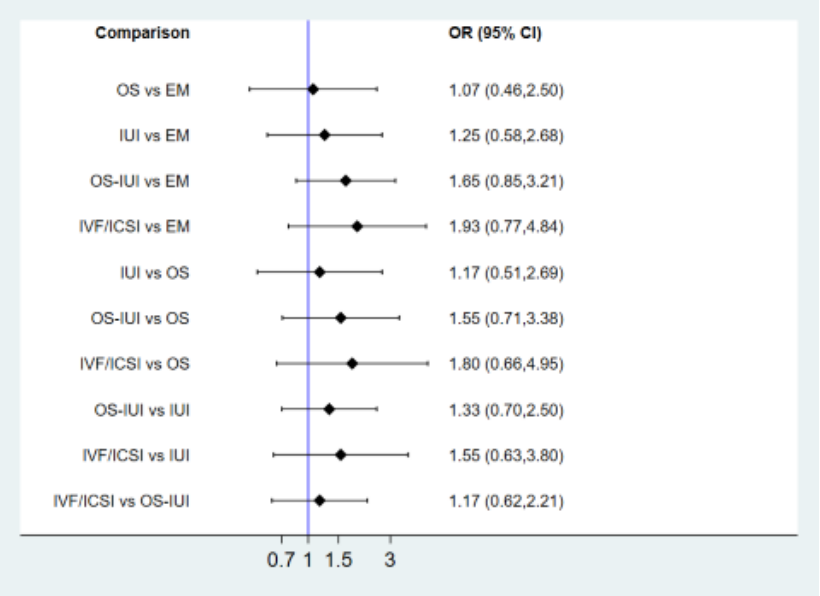


Figure 12 Sensitivity analysis for live birth by exclusion of abstract-only publications.

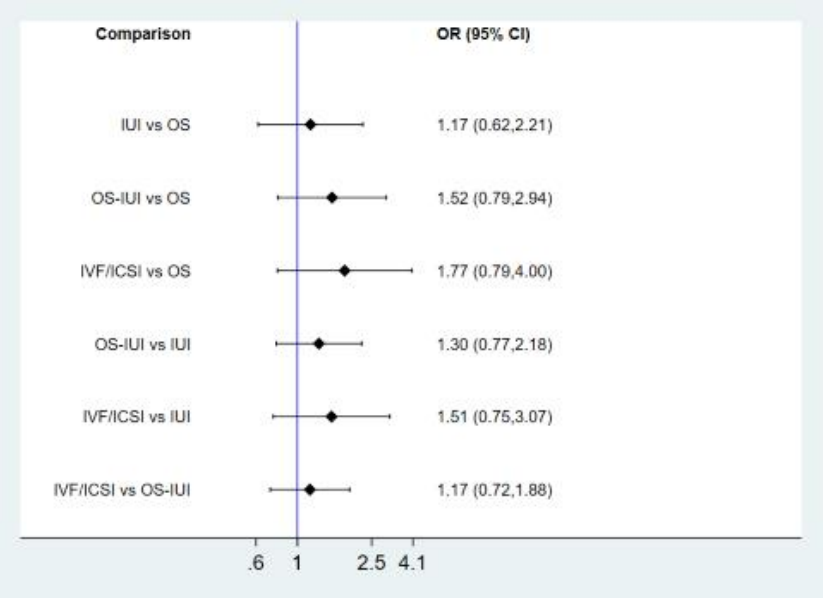


Figure 13 Sensitivity analysis for live birth by excluding RCTs involving expectant management from the network

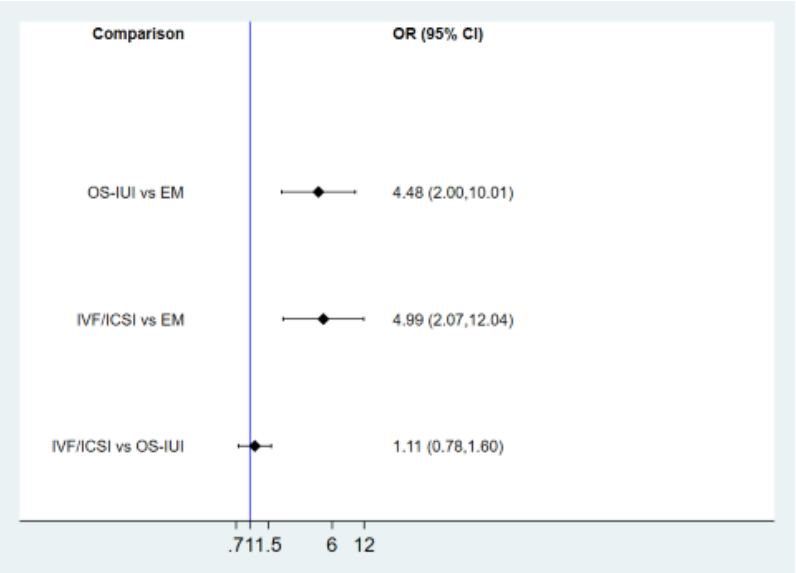


Figure 14 Sensitivity analysis for live birth by limiting to RCTs on couples with poor prognosis of natural conception.

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.

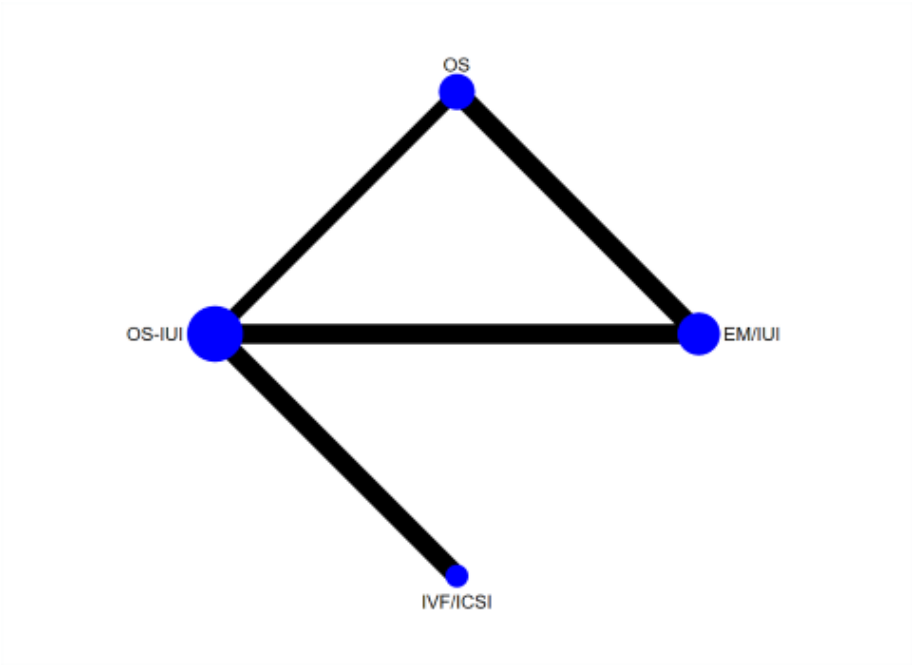


Figure 15 Network plot for multiple pregnancy.

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).

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 (Appendix 7.2) are consistent with those in the network meta-analysis.

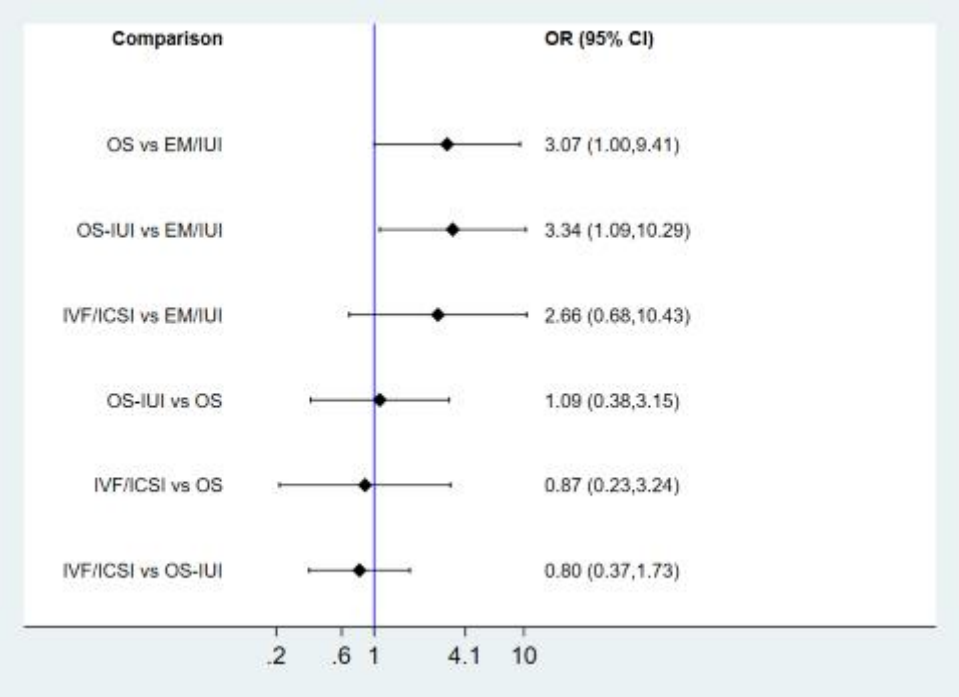


Figure 16 Network meta-analysis for multiple pregnancy.

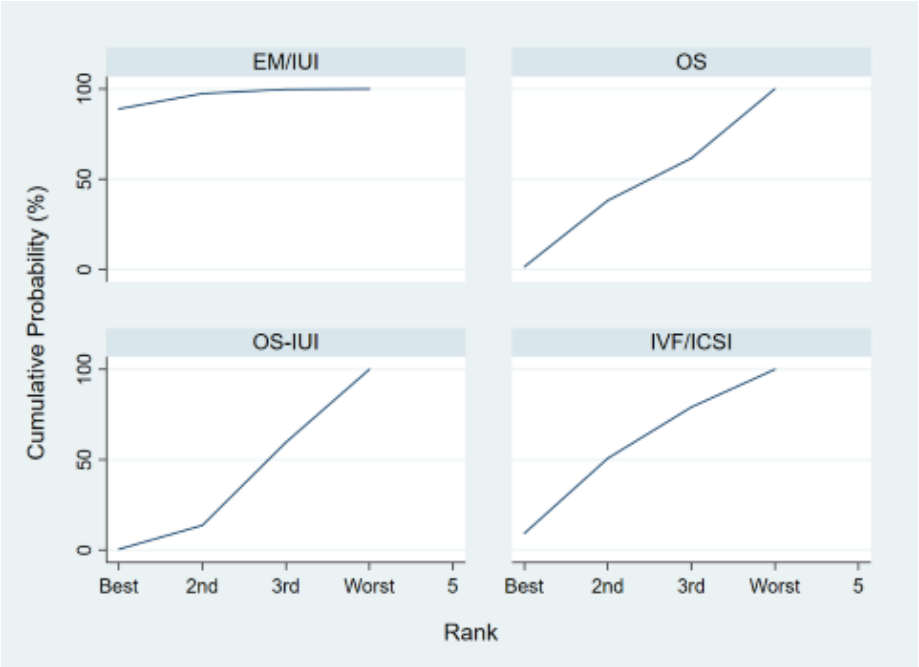


Figure 17 Cumulative rankograms of interventions for multiple pregnancy.

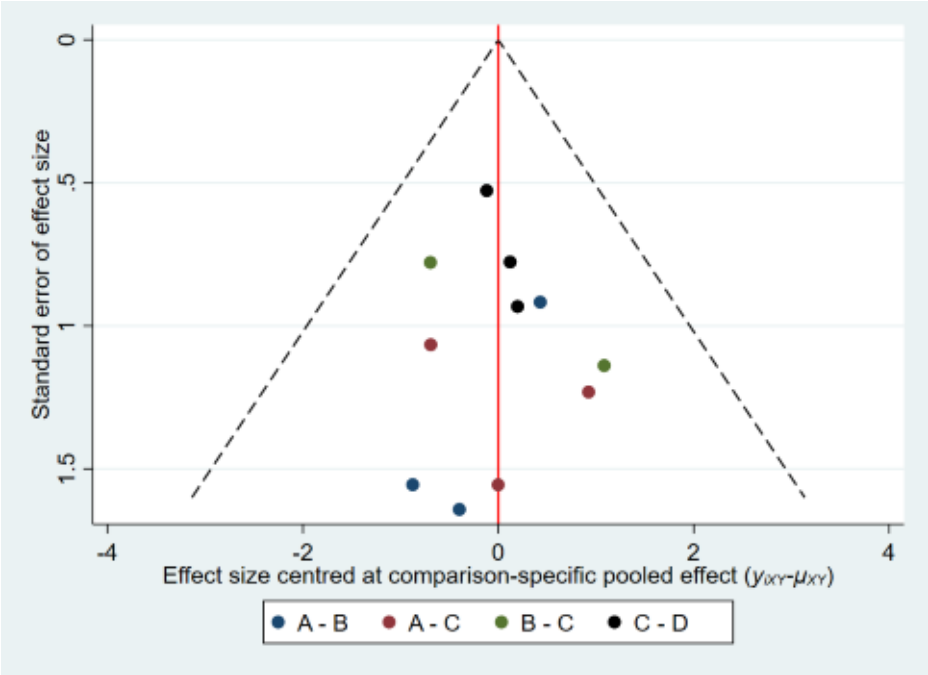


Figure 18 Comparison-adjusted funnel plot for multiple pregnancy.
(A: expectant management or IUI; B: OS; C: OS-IUI; D: IVF/ICSI.)

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., 1991b; 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., 2014a; 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.

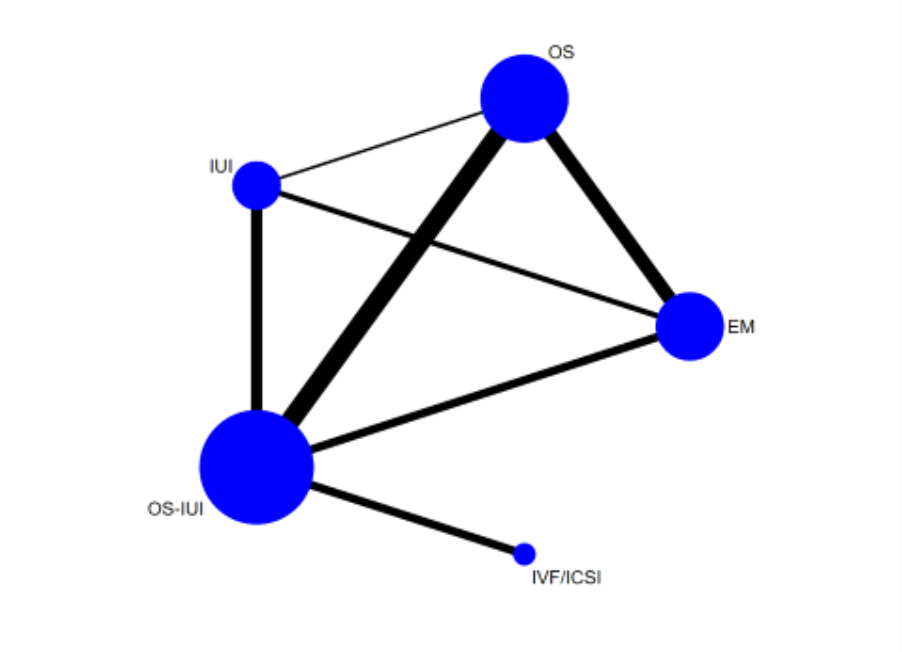


Figure 19 Network plot for clinical pregnancy.

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 (Appendix 7.3).

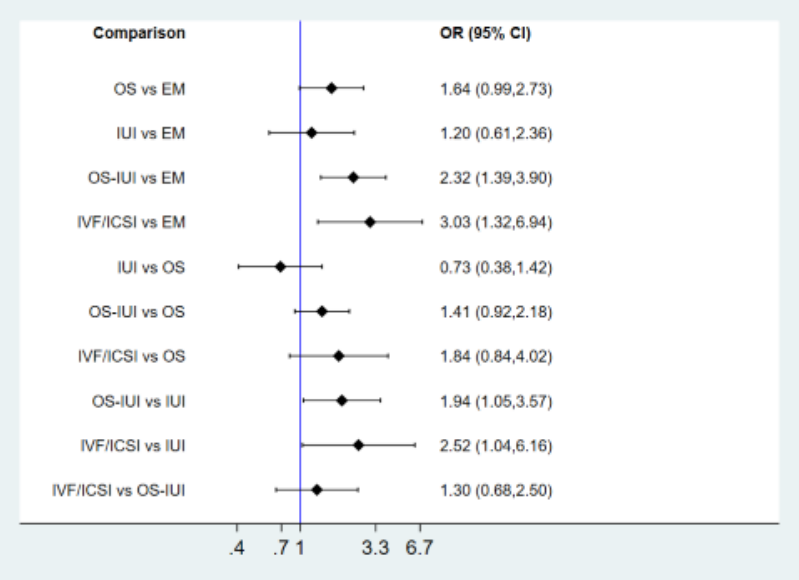


Figure 20 Network meta-analysis for clinical pregnancy.

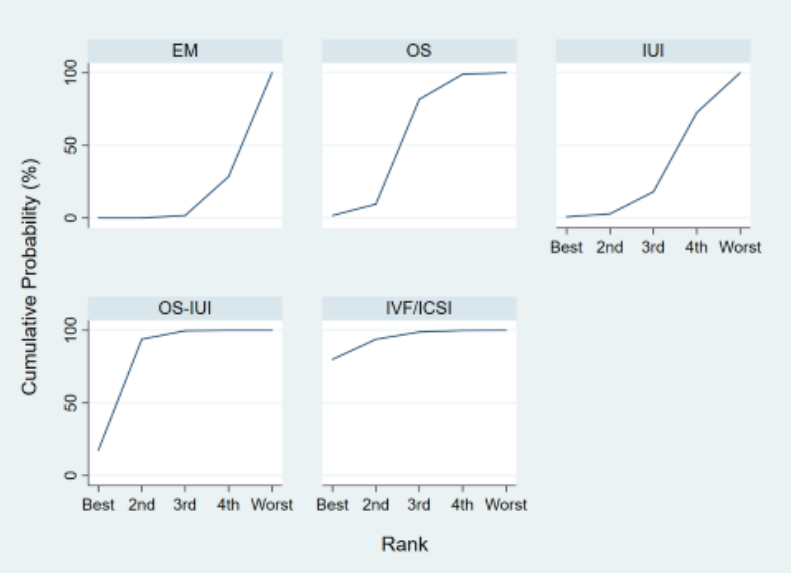


Figure 21 Cumulative rankograms of interventions for clinical pregnancy.

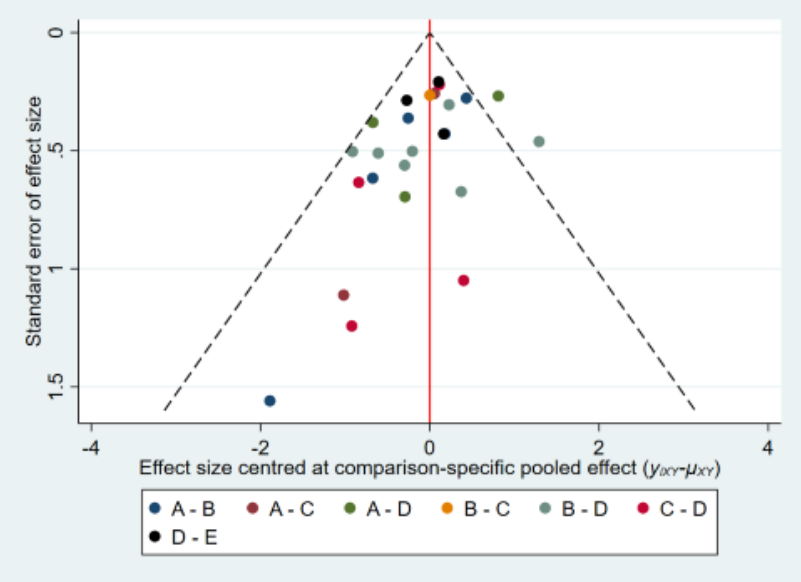


Figure 22 Comparison-adjusted funnel plot for clinical pregnancy.
(A: expectant management; B: OS; C: IUI; D: OS-IUI; E: IVF/ICSI.)

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%.

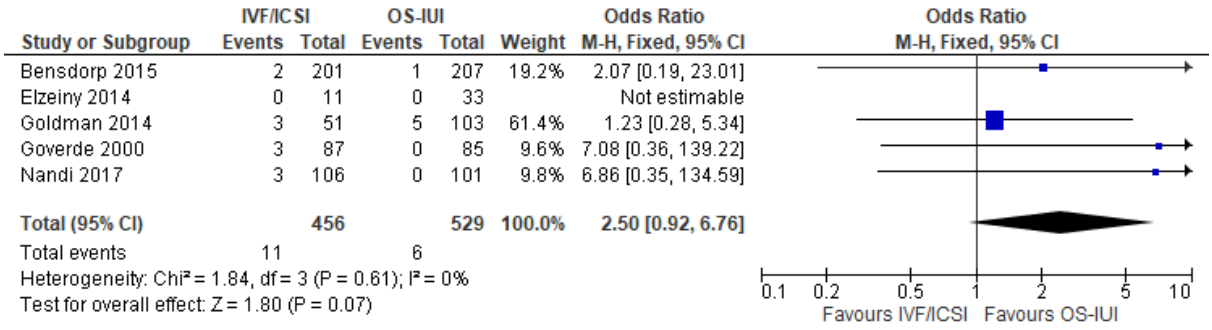


Figure 23 Forest plot of comparison: 2 Pairwise meta-analysis for OHSS, outcome: 2.5 IVF/ICSI vs OS-IUI.

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.

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).

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.

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* 1991a;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* 1991b;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* 2014a;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* 2014b;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.

Chapter 6

Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis

This is the accepted version of the following article: Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, Mijatovic V, Mol BW. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. *Ultrasound Obstet Gynecol*. 2019; 54(2):172-181, which has been published in final form at: https://obgyn.onlinelibrary.wiley.com/doi/full/10.1002/uog.20238

Statement of Authorship

Title of Paper	Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wang R, van Welie N, van Rijswijk J, Johnson NP, Norman RJ, Dreyer K, Mijatovic V, Mol BW. Effectiveness on fertility outcome of tubal flushing with different contrast media: systematic review and network meta-analysis. Ultrasound Obstet Gynecol. 2019; 54(2):172-181

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	RW conceptualised and designed the study, collected the data, analysed the data, interpreted the data, commented on the drafts, approved the final draft and acted as the corresponding author.		
Overall percentage (%)	75%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	2/8/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

i. the candidate's stated contribution to the publication is accurate (as detailed above);

ii. permission is granted for the candidate in include the publication in the thesis; and

iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Nienke van Welie		
Contribution to the Paper	NvW collected the data, interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

Name of Co-Author	Joukje van Rijswijk
Contribution to the Paper	JvR interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.

Signature		Date	2/8/2019

Name of Co-Author	Neil P Johnson		
Contribution to the Paper	NPJ interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

Name of Co-Author	Robert J Norman		
Contribution to the Paper	RJN interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

Name of Co-Author	Kim Dreyer		
Contribution to the Paper	KD interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

Name of Co-Author	Velja Mijatovic		
Contribution to the Paper	VM interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

Name of Co-Author	Ben Willem J Mol		
Contribution to the Paper	BWM conceptualised and designed the study, interpreted the data, revised the manuscript for important intellectual content, commented on the drafts and approved the final draft.		
Signature		Date	2/8/2019

ABSTRACT

Objectives: To compare, in women with infertility, the effectiveness and safety of tubal flushing using oil-based contrast medium, water-based contrast medium or their combination, and no tubal flushing, and to evaluate the effectiveness of tubal flushing on fertility outcome over time.

Methods: We performed a systematic review and network meta-analysis, searching the electronic databases MEDLINE, EMBASE and Cochrane Central Register of Controlled Trials, and trial registries, up to 25 September 2018. We included randomized controlled trials (RCTs) comparing the following interventions with each other or with no intervention in women with infertility: tubal flushing using water-based contrast medium, tubal flushing using oil-based contrast medium or additional tubal flushing with oil-based medium following diagnostic tubal flushing with water-based medium. The outcomes included clinical pregnancy, live birth, ongoing pregnancy, miscarriage, ectopic pregnancy and adverse events.

Results: Of the 283 studies identified through the search, 14 RCTs reporting on 3852 women with infertility were included. Network meta-analysis showed that tubal flushing using oil‐based contrast medium was associated with higher odds of clinical pregnancy within 6 months after randomization and more subsequent live births compared with tubal flushing using water-based medium (odds ratio (OR) 1.67, 95% CI 1.38–2.03, moderate certainty of evidence; and OR 2.18, 95% CI 1.30–3.65, low certainty of evidence, respectively) and compared with no intervention (OR 2.28, 95% CI 1.50‐3.47, moderate certainty of evidence; and OR 2.85, 95% CI 1.41-5.74, low certainty of evidence, respectively). These results agreed with those of the pairwise meta-analysis. For clinical pregnancy within 6 months, there was insufficient evidence of a difference between tubal flushing with water-based contrast medium and no intervention (OR 1.36, 95% CI 0.91-2.04, low certainty of evidence). For fertility outcomes after 6 months, there was insufficient evidence of a difference in any comparison (low to very low certainty of evidence). Compared with tubal

flushing using water‐based contrast medium, the use of oil‐based contrast medium was associated with higher odds of asymptomatic intravasation (OR 5.06, 95% CI 2.29–11.18, moderate certainty of evidence).

Conclusion: In women with infertility undergoing fertility workup, tubal flushing using oil‐based contrast medium probably increases clinical pregnancy rates within 6 months after randomization and may increase subsequent live‐birth rates, compared with tubal flushing using water‐based contrast medium and compared with no intervention. Evidence on fertility outcomes beyond 6 months is inadequate to draw firm conclusions.

PROSPERO Registration: CRD42017059832

Keywords: fallopian tube patency tests, tubal flushing, hysterosalpingography, HyCoSy, laparoscopy, contrast media, infertility, systematic review

INTRODUCTION

Tubal flushing was initially introduced in reproductive medicine as a diagnostic test to evaluate tubal patency. It constitutes an essential part of the fertility work-up, as recommended in clinical guidelines (National Institute for Health and Care Excellence, 2013; Practice Committee of the American Society for Reproductive Medicine, 2015). It has been used in several different techniques to visualise tubal patency, including hysterosalpingography (HSG), hysterosalpingo-contrast sonography (HyCoSy), hysterosalpingo-foam sonography (HyFoSy) and laparoscopy with dye testing. Water-based contrast has been widely used in all these procedures and oil-based contrasts are mainly used in HSG.

Debates about the therapeutic effects of tubal flushing started over six decades ago (King and Herring, 1949; Weir and Weir, 1951). Several potential mechanisms have been proposed to explain such therapeutic effects, including mechanical flushing out the debris or mucus plugs in the Fallopian tubes (Gillespie, 1965), enhancement of ciliary activity (Soules and Spadoni, 1982) and immunobiological actions on the endometrium or peritoneum (Izumi et al., 2017; Johnson, 2014; Mikulska et al., 1994; Sawatari et al., 1993; Yun and Lee, 2004).

In order to evaluate the effects of tubal flushing on fertility outcomes, a number of studies compared tubal flushing with different contrast media, alone or in combination, with each other or no treatment. However, no large RCTs have compared all these different interventions, and therefore a network meta-analysis incorporating both direct and indirect evidence is required to determine the most effective contrast media in imaging techniques. Moreover, based on the available potential mechanisms, the effectiveness of tubal flushing may not be the same over time, and therefore it is also important to assess the trend of fertility outcomes with different contrast media over time.

Several meta-analyses on this topic have been published (Fang et al., 2018; Mohiyiddeen et al., 2015; Watson et al., 1994). These meta-analyses only used direct evidence in the evidence synthesis and some evaluated water- versus oil-based contrast in HSG only

(Mohiyiddeen et al., 2015; Watson et al., 1994), but did not consider women without tubal flushing or women with tubal flushing undergoing non-HSG techniques. None of these meta-analyses has consider the fertility outcomes over time (Fang et al., 2018; Mohiyiddeen et al., 2015; Watson et al., 1994).

We therefore conducted this systematic review and network meta-analysis to compare the effectiveness and safety of tubal flushing with oil- or water-based contrast, alone or in combination, with each other or with no tubal flushing in women with infertility undergoing fertility work-up. Our secondary objective was to evaluate the effectiveness of tubal flushing on fertility outcomes over time.

MATERIALS AND METHODS

The protocol of this systematic review was registered on PROSPERO (CRD42017059832). We reported the systematic review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) extension statement for network meta-analysis (Hutton et al., 2015).

Information sources and search strategies

We searched the electronic database including EMBASE, Cochrane Central Register of Controlled Trials (CENTRAL) and MEDLINE as well as the trial registers (ClinicalTrials.gov, International Clinical Trials Registry Platform and Australian New Zealand Clinical Trials Registry) and reference lists of identified publications. The last electronic database search was conducted on 25th September 2018 (Appendix S1).

Eligibility criteria

The study population intended to include all women wishing to conceive. We included RCTs comparing at least two of the following treatment or control groups: (1) no tubal flushing; (2) tubal flushing with water-based contrast; (3) tubal flushing with oil-based contrast; or (4) an additional tubal flushing procedure with oil-based contrast after diagnostic tubal flushing with water-based contrast. Tubal flushing procedures using any imaging technique including HSG, HyCoSy, HyFoSy or laparoscopy (or hydrolaparoscopy) was eligible. Studies compared different types of water-based contrast (or oil-based contrast) were excluded. Quasi-RCTs were excluded. No language limitation was applied.

Outcomes

The outcomes included clinical pregnancy, live birth, ongoing pregnancy, miscarriage, ectopic pregnancy and adverse events. We intended to use outcomes at the longest time of follow-up in each study as the primary analysis. In order to show the trend over time, we

also planned subgroup analysis to evaluate clinical pregnancy at different follow-up time points (i.e. 3, 6, 9, 12 and 18 months) after randomisation if data were available. All short-term outcomes related to tubal flushing such as pelvic infection and intravasation as well as long-term outcomes like birth defects were reported on.

Study selection, data collection and quality assessment

Two reviewers (RW and NvW) independently evaluated study eligibility, extracted the data and assessed the quality of included studies. Disagreements were solved by consensus or by a third reviewer (BWJM).

We used a predesigned form to collect the following information: name of the first author, publication year, study population, participants' characteristics, funding, types of contrast media, details of interventions and co-interventions, sample sizes and outcomes. If outcome data were available in published curves, we used DigitizeIt 2.2 software to reconstruct the data from publications (Guyot et al., 2012).

We assessed risk of bias within individual studies by using the Cochrane Collaboration's tool (Higgins, 2011) and evaluated the certainty across studies in study limitations (risk of bias), indirectness, inconsistency, imprecision, and publication bias by using Confidence in Network Meta-analysis (CINeMA) (Salanti et al., 2014).

Statistical analysis

We used network plots to show available head-to-head comparisons in included RCTs and used contribution matrix to illustrate the contribution of each head-to-head comparison to the overall body of evidence (Chaimani et al., 2013; Chaimani and Salanti, 2015).

We then tested global inconsistency by using the design-by-treatment interaction model (Higgins et al., 2012) and tested local inconsistency by using inconsistency plots (Dias et al., 2010). When there was no significant inconsistency, we performed network meta-analyses

within multivariate random effects meta-analysis models (White, 2015) as well as random effects pairwise meta-analysis (Higgins, 2011).

We used the surface under the cumulative ranking (SUCRA) to rank the treatments (Salanti et al., 2011) and apply the comparison adjusted funnel plot to assess small study effects (Chaimani et al., 2013). We used STATA (version 15.0, StataCorp) to perform statistical analysis and to illustrate the graphics (White, 2015).

We intended to perform subgroup analyses on age, duration of infertility, cause of infertility and outcomes at different time points if the data were available. We also planned a sensitivity analysis by including only studies with low risk of bias. We performed a post-hoc sensitivity analysis by excluding the participants with missing outcome data (Higgins, 2011).

RESULTS

Characteristics of included studies

Of the 283 studies identified, 14 RCTs (16 articles) reporting on 3,852 women with infertility were included (Al-Fadhli et al., 2006; Alper et al., 1986; de Boer et al., 1988; Dreyer et al., 2017; Johnson et al., 2004; Johnson et al., 2007; Letterie and Rose, 1990; Lindborg et al., 2009; Lindequist et al., 1991; Lindequist et al., 1994; Nugent et al., 2002; Ogata et al., 1993; Rasmussen et al., 1991; Spring et al., 2000; Steiner et al., 2003; Yang et al., 1989) (Figure 1). All studies reported on women with infertility and at least included unexplained infertility, while three (Lindequist et al., 1994; Ogata et al., 1993; Rasmussen et al., 1991) did not report the detailed causes of infertility (Table 1). Funding was reported in three studies (Dreyer et al., 2017; Johnson et al., 2004; Lindborg et al., 2009). Outcome data at different time points were extracted from the graphics in seven studies by using DigitizeIt 2.2 software (Al-Fadhli et al., 2006; Dreyer et al., 2017; Johnson et al., 2007; Lindborg et al., 2009; Lindequist et al., 1994; Rasmussen et al., 1991; Steiner et al., 2003).

Of the 14 included studies, the most frequent comparison was water- versus oil-based contrast (n=6) (Alper et al., 1986; de Boer et al., 1988; Dreyer et al., 2017; Lindequist et al., 1994; Rasmussen et al., 1991; Spring et al., 2000), followed by both versus water-based contrast (n=5) (Al-Fadhli et al., 2006; Letterie and Rose, 1990; Spring et al., 2000; Steiner et al., 2003; Yang et al., 1989). There were three studies comparing oil-based contrast to no tubal flushing (Johnson et al., 2004; Nugent et al., 2002; Ogata et al., 1993), one comparing both to oil-based contrast (Spring et al., 2000), and one comparing water-based contrast to no tubal flushing (Lindborg et al., 2009) (Table 1). Clinical pregnancy within 6-month was the most commonly reported outcome (n=12). The network plots for different outcomes were presented in Figure S1.

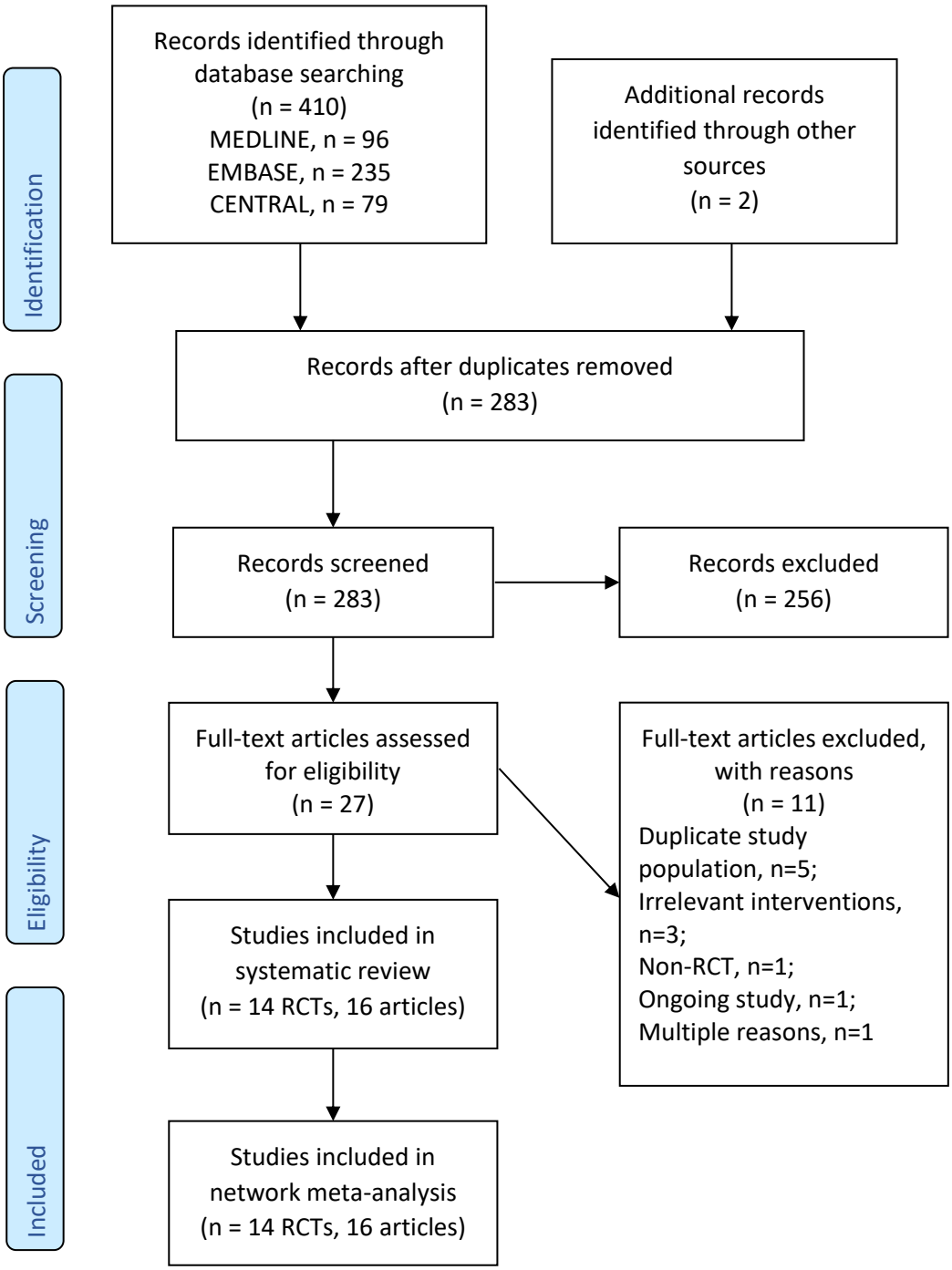


Figure 1 PRISMA flow diagram.

Table 1 Characteristics of included studies

Study	Country	Causes of infertility	Sample size	Follow-up months	Intervention and control	Method of tubal flushing	Contrast media	Co-intervention during/after tubal flushing
Al-Fadhli 2006	Canada	Unexplained infertility, endometriosis	88	12	Water+oil	Laparoscopy	1) Dilute solution of methylene blue dye; 2) Lipiodol	Excision of the endometriotic lesions when necessary
					Water	Laparoscopy	1) Dilute solution of methylene blue dye; 2) Saline 10 ml.	Excision of the endometriotic lesions when necessary
Alper 1986	Canada/USA	Oligo-ovulation (33%), tubal factor (29%), endometriosis (8%), Unexplained or mild male infertility (30%)	131	6	Oil	HSG	Lipiodol 10-20 ml	Ovulation disorders were treated with clomiphene citrate.
					Water	HSG	Reno-M-60 (diatrizoate) 10-20ml	Ovulation disorders were treated with clomiphene citrate.
De Boer 1988	Netherlands	Unexplained or mild male infertility	175	6	Oil	HSG	Ethiodol 10 ml	"Only hormonal therapy was used".
					Water	HSG	Iopamidol 10 ml	"Only hormonal therapy was used".
Dreyer 2017	Netherlands	Unexplained or mild male infertility (87%), tubal factor (8%), other causes (5%). all had low risk of tubal pathologies	1119	6	Oil	HSG	Lipiodol 5-10 ml	According to prespecified indications: IUI (17.9%), IVF/ICSI (1.4%), ovulation induction (0.5%), laparoscopy (6.2%) or hysteroscopy (4.4%)
					Water	HSG	Telebrix Hystero (Meglumine Ioxitalamate) 5-10ml	According to prespecified indications: IUI (14.8%), IVF/ICSI (1.3%), ovulation induction (0.5%),

								laparoscopy (6.2%) or hysteroscopy (4.2%)
Johnson 2004	New Zealand	Unexplained infertility (61%), endometriosis (39%)	158	6	Oil	HSG	Lipiodol 10 ml	No other co-interventions
					No flushing	NA	NA	No other co-interventions
Letterie 1990	USA	Unexplained infertility (100%)	40	12	Water+oil	Laparoscopy	1) Laparoscopic tubal testing; 2) Ethiodol 20 ml	No other co-interventions
					Water	Laparoscopy	1) Laparoscopic tubal testing; 2) Conray-60 20 ml	No other co-interventions
Lindborg 2009	Sweden	Unexplained or mild male infertility (77%), tubal factor (23%)	334	6	Water	HyCoSy	Echovist <15 ml	No other co-interventions (but 4 women had IVF)
					No flushing	NA	NA	No other co-interventions
Lindequist 1994	Denmark	NA	242	9	Oil	HSG	Lipiodol 5-10 ml	No other co-interventions
					Water	HSG	Iotrolan (Isovist) 5-10ml	No other co-interventions
Nugent 2002	UK	Unexplained infertility	34	6	Oil	HSG	Lipiodol 5.8ml (mean)	No other co-interventions
					No flushing	NA	NA	No other co-interventions
Ogata 1993	Japan	NA	302	4	Oil	HSG	Lipiodol, volume not reported	No other co-interventions
					No flushing	NA	NA	No other co-interventions
Rasmussen 1991	Denmark	Tubal factor (46%) and unexplained or mild male infertility (54%)	398	10	Oil	HSG	Lipiodol 5-10ml	No other co-interventions
					Water	HSG	Iohexol, Ioxaglate or Diatrizoate 5-10ml	No other co-interventions

Spring 2000	USA	Tubal, male, ovulatory, age, uterine, peritoneal, endocrine, cervical, immunologic, drug-associated factor or unexplained infertility[c]	666	12	Water+oil	HSG	1) Sinografin (52.7% diatrizoate meglumine and 26.8% iodipamide meglumine) 8.2 ml (mean); 2) Ethiodol 6.0 ml (mean)	IUI (24.8%)
					Oil	HSG	Ethiodol 8.6 ml (meal)	IUI (25.3%)
					Water	HSG	Sinografin (52.7% diatrizoate meglumine and 26.8% iodipamide meglumine) 9.4 ml (mean)	IUI (24.8%)
Steiner 2003	USA	Ovulatory dysfunction (35%), endometriosis (12%), unexplained or mild male infertility (45%), tubal factor (3%)	56	18	Water+oil	HSG	1) Sinografin (52.7% diatrizoate meglumine and 26.8% iodipamide meglumine) 5–10 mL; 2) Ethiodol 10 ml	Ovulatory medications (53.3%)
					Water	HSG	Sinografin (52.7% diatrizoate meglumine and 26.8% iodipamide meglumine) 5–10 ml	Ovulatory medications (61.5%)
Yang 1989	China	Unexplained or mild male infertility (45%), anovulation (23%), endometriosis (8%) Other factor (24%)	109	8	Water+oil	HSG	1) Telebrix hystero 10ml; 2) Lipiodol 5 ml	No other co-interventions
					Water	HSG	Telebrix hystero 10ml	No other co-interventions

Footnotes: NA: not applicable; a. median of the control group; b. oil: oil-based contrast; water: water-based contrast; none: no tubal flushing; both: tubal flushing with water-based contrast followed by oil-based contrast; c. detailed data were not available due to incomplete report of the baseline data.

Quality of evidence of individual studies

With regard to selection bias, 64% of included RCTs (n=9) reported adequate methods of random sequence generation (Al-Fadhli et al., 2006; Alper et al., 1986; Dreyer et al., 2017; Johnson et al., 2004; Letterie and Rose, 1990; Lindborg et al., 2009; Spring et al., 2000; Steiner et al., 2003; Yang et al., 1989), and 36% (n=5) reported adequate methods of allocation concealment (Dreyer et al., 2017; Johnson et al., 2004; Lindborg et al., 2009; Nugent et al., 2002; Yang et al., 1989) while 7% (n=1) had no concealment (Steiner et al., 2003) (Table 2). As blinding was not possible due to the nature of the interventions, we scored the risk of performance bias as unclear in all RCTs. Given that all the fertility outcomes are objective outcomes, it is unlikely that the non-blinded design will affect the outcome measurement and therefore the risk of detection bias was low in all the included studies. Five RCTs (36%) had high risk of attrition bias due to the considerable proportion of missing outcome data (Al-Fadhli et al., 2006; Alper et al., 1986; Letterie and Rose, 1990; Lindequist et al., 1994; Ogata et al., 1993). One RCT (7%) was scored at high risk of other bias as the age distribution was imbalanced in groups (Spring et al., 2000).

The majority body of evidence (>70%) in comparisons between water vs none, water vs none and oil vs water are at low risk of bias, but the evidence in the comparisons between the combination group and others are prone to be biased as at least 25% of the evidence was at high risk of bias (Figure S3).

Network consistency and contribution

When considering clinical pregnancy at the longest time of follow-up in each study in the analysis, we found significant global and local inconsistency (Table S1 and Figure S3). Therefore, pooling the outcomes at different time points should be avoided (Higgins et al., 2012). We chose clinical pregnancy at the most commonly used time point (6-month) as an alternative outcome and presented outcomes at other time points in subgroup analyses. After separating outcomes at different time points, no significant global or local inconsistency was

observed (Table S1 and Figure S3). Therefore, time of outcome measurement is an important source of inconsistency in this network meta-analysis. The contribution of direct evidence to the network for different outcomes were presented in Figure S4.

Table 2 Risk of bias of included studies

Study	Selection bias (random sequence generation)	Selection bias (allocation concealment)	Performance bias (blinding of participants and [illegible])	Detection bias (blinding of outcome assessment)	Attrition bias (incomplete outcome data)	Re0orting bias (selective reporting)	Other bias (other sources of bias)
Al-Fadhli 2006	+	?	?	+	-	?	?
Alper 1986	+	?	?	+	-	?	?
De Boer 1988	?	?	?	+	+	?	+
Dreyer 2017	+	+	?	+	+	+	+
Johnson 2004	+	+	?	+	+	+	+
Letterie 1990	+	?	?	+	-	?	?
Lindborg 2009	+	+	?	+	+	+	+
Lindequist 1994	?	?	?	+	-	?	+
Nugent 2002	?	+	?	+	+	+	+
Ogata 1993	?	?	?	+	-	?	+
Rasmussen 1991	?	?	?	+	?	?	?
Spring 2000	+	?	?	+	+	+	-
Steiner 2003	+	-	?	+	+	?	+
Yang 1989	+	+	?	+	+	?	?

Footnote: "+" low risk of bias; "?" unclear risk of bias; "-" high risk of bias.

Network and pairwise meta-analyses

Clinical pregnancy

Twelve RCTs reported clinical pregnancy within 6 months in 2,884 women. Network meta-analysis (Figure 2) showed oil-based contrast increased the odds of clinical pregnancy compared to no tubal flushing (OR 2.28, 95% CI 1.50-3.47, moderate certainty of evidence), while there was insufficient evidence of a difference between water-based contrast and no tubal flushing (OR 1.36, 95% CI 0.91-2.04, low certainty of evidence). This suggests that if the 6-month clinical pregnancy rate following no tubal flushing is assumed to be 16%, the clinical pregnancy rate following tubal flushing with oil-based contrast and water-based contrast would be 30% (22%-40%) and 21% (15%-28%), respectively. Compared to water-based contrast, oil-based contrast resulted in higher odds of clinical pregnancies (OR 1.67, 95% CI 1.38-2.03, moderate certainty of evidence). This suggests that if the 6-month clinical pregnancy rate following tubal flushing with water-based contrast is assumed to be 28%, the clinical pregnancy rate following tubal flushing with oil-based contrast would be 39% (35%-44%). The evidence on the comparison between an additional oil-based tubal flushing to a water-based tubal flushing versus the other interventions were at very low certainty of evidence. SUCRA values for the combination, oil-, water-based contrast and no tubal flushing were 83.0%, 82%, 31.7% and 2.5%, respectively. Pairwise meta-analysis showed similar results in these comparisons (Figure 2). There was no evidence of existence of small study effects (Figure S5).

Subgroup analysis of clinical pregnancy within 3 months showed similar results (Figure S6). With regard to clinical pregnancy within 9, 12 and 18 months, network meta-analyses nor pairwise meta-analyses showed any statistically significant differences in most comparisons. As the breakdown outcome data on women with different ages, durations of infertility, or causes of infertility were not available, no subgroup analyses on these variables were performed. Sensitivity analyses on studies with overall low risk of bias and after excluding participants with missing outcome data showed consistent results (Figure S7).

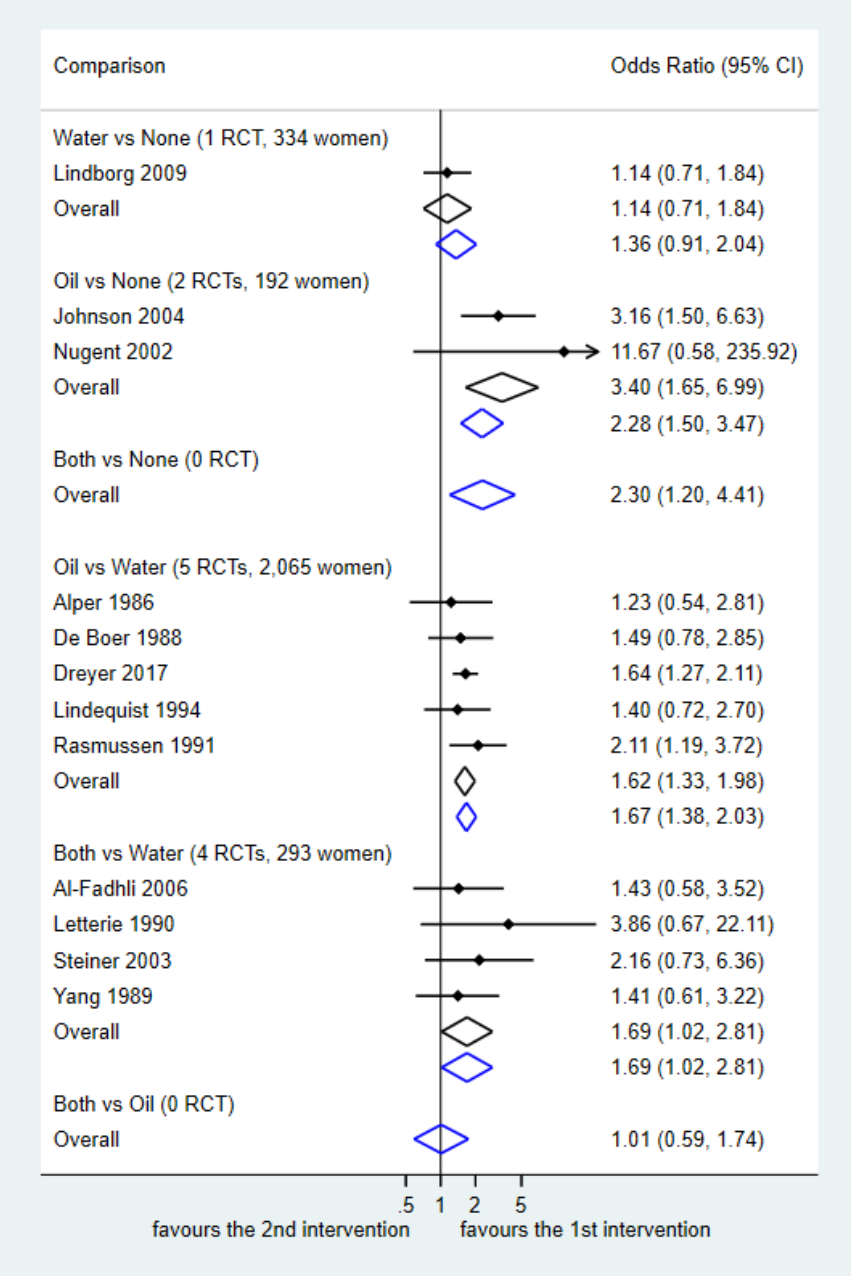


Figure 2 Forest plot of network and pairwise meta-analyses on clinical pregnancy
Clinical pregnancies within 6 months are presented in this forest plot. ORs and 95% CIs of pairwise meta-analyses are illustrated in black diamonds while those of network meta-analyses are illustrated in blue diamonds.

Live birth and ongoing pregnancy

Five studies reported on live birth resulting from pregnancy within 6 months in 2,043 women. Network meta-analysis (Figure 3) showed that oil-based contrast resulted in higher odds of live birth compared to no tubal flushing (OR 2.85, 95% CI 1.41-5.74, low certainty of evidence), while there was insufficient evidence of a difference between water-based and no tubal flushing (OR 1.31, 95% CI 0.70-2.44, certainty of evidence). This suggests that if the live birth rate following no tubal flushing is assumed to be 16%, the live birth rate

following tubal flushing with oil-based contrast and water-based contrast would be 35% (21%-52%) and 20% (12%-32%), respectively. Oil-based contrast resulted in higher odds of live birth compared to water-based contrast (OR 2.18, 95% CI 1.30-3.65, low certainty of evidence). This suggests that if the live birth rate following tubal flushing with water-based contrast is assumed to be 22%, the live birth rate following tubal flushing with oil-based contrast would be 38% (27%-51%). SUCRA values for oil-, water-based contrast and no tubal flushing were 99.9%, 41.1% and 9.0%, respectively. Results of pairwise meta-analyses were consistent with those in network meta-analysis (Figure 3). The results for 6-month ongoing pregnancy (4 RCTs, 1,645 women) were consistent with those for live birth (Figure S6).

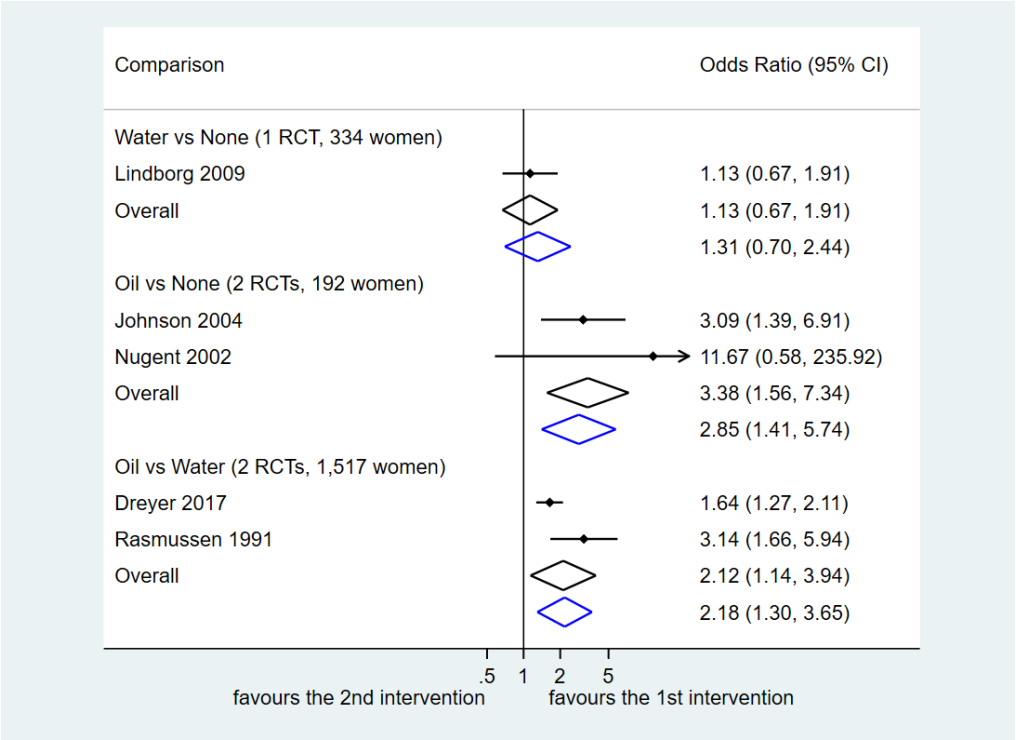


Figure 3 Forest plot of network and pairwise meta-analyses on live birth
Live births resulting from pregnancy within 6 months are presented in this forest plot. ORs and 95% CIs of pairwise meta-analyses are illustrated in black diamonds while those of network meta-analyses are illustrated in blue diamonds.

Miscarriage, ectopic pregnancy and adverse events

There was no conclusive evidence of a difference in any of the comparisons for miscarriage or ectopic pregnancy (Figure S6).

Five studies reported no short-term adverse events. Pooled analysis of the three studies showed that, compared to water-based contrast, oil-based contrast was associated with higher odds of asymptomatic intravasation (OR 5.06, 95% CI 2.29-11.18, 3 studies, I^2=0, Figure S6). No case of pulmonary embolism or death was reported. Three studies reported pelvic infection, in which two studies compared water- and oil-based contrast. Pool analysis showed that there was insufficient evidence of differences in pelvic infection between these two interventions (OR 0.23, 95% CI 0.04-1.27, 2 studies, I^2=0, Figure S6). Only one study reported long-term adverse events (Dreyer et al., 2017), in which three new-borns in the oil-based contrast group had skeletal dysplasia, oesophageal atresia, and chromosomal mosaicism, respectively. No congenital abnormalities were seen in the water group.

DISCUSSION

Summary of key findings

For the comparisons between oil-based contrasts versus water-based contrast or no tubal flushing, the overall certainty of evidence was moderate for short-term clinical pregnancy, low for short-term live birth and low to very low for outcomes beyond 6 months. Tubal flushing with oil-based contrast probably increases short-term (6 months) clinical pregnancy rate and may increase subsequent live birth rate compared to tubal flushing with water-based contrast and no tubal flushing, but it is not certain whether such potential superiority persists beyond 6 months. There evidence on the effectiveness of water-based contrast was insufficient (low certainty of evidence).

Strengths and limitations

The strengths of this systematic review included incorporating both evidence from direct and indirect comparisons, providing hierarchy of rankings of effectiveness and using multiple approaches for sensitivity analyses. Moreover, time was incorporated into the outcome assessments and was identified as a source of inconsistency. Our current outcome reporting strategy not only reduced the heterogeneity in outcome reporting across trials but also illustrated the trend of effectiveness of different interventions over time.

Meanwhile, several limitations of our meta-analysis should be addressed. Firstly, not all trials reported live birth. The agreements between the results for live birth and clinical pregnancy gave some reassurance for the conclusion. Secondly, although the majority of the participants were broadly defined as unexplained infertility, including endometriosis and mild male factor infertility, the study population also included other different causes of infertility. The heterogeneous nature of the study population may result in selection bias. Thirdly, some of the studies had high risk of selection, attrition or other bias. This resulted in overall low quality of evidence in some comparisons, especially in the combination group. Finally, the association between available evidence and competing interests of the

manufacturer was unclear in most studies. Only three (Dreyer et al., 2017; Johnson et al., 2004; Spring et al., 2000) reported funding sources and all were from academic institutes or societies, including one (Johnson et al., 2004) with additional support from the industry.

Clinical implications

The effect of oil-based contrast has not been evaluated outside HSG. HyCoSy is an accurate test for diagnosing tubal occlusion and performs similarly to HSG (Maheux-Lacroix et al., 2014). It prevents women from exposure to radiation and therefore has replaced HSG in fertility workup in many settings. More recently, hysterosalpingo-foam sonography (HyFoSy) also became commonly used and a trial on its diagnostic accuracy and cost-effectiveness in fertility work-up is underway (van Rijswijk et al., 2018).

The costs of HyCoSy with sonographic contrast and HSG with water-based contrast are considered similar (Lim et al., 2011) and HyFoSy may further reduce the cost (van Rijswijk et al., 2018). In HSG, oil-based contrast is more expensive than water-based contrast, with an extra US$8,198 for an additional ongoing pregnancy in ovulatory women with infertility at low risk for tubal pathology (van Rijswijk et al., 2018). Cost-effectiveness should also be considered in shared-decision making.

Several safety concerns on tubal flushing have been raised. Firstly, venous intravasation occurs in approximately 2-7% cases in HSG (Bateman et al., 1980; Dusak et al., 2013; Nunley et al., 1987) and occurs more frequently when using oil-based contrast. Some reports on venous intravasation using ultrasound show a higher incidence (13%) (Wang et al., 2018). While intravasation can potentially result in life threating pulmonary embolism, we are unaware of any deaths reported since the 1960s (Siegler, 1967). This may be due to fluoroscopy screening or the reduced use of HSG with oil-based contrast worldwide as HyCoSy and HyFoSy become more popular. Secondly, the concern about the thyroid function of mother and child is based on the effects of iodinated contrast media (Satoh et al., 2015; So et al., 2017) and a longer persisting time of oil-based contrast in the pelvis

(Miyamoto et al., 1995). Maternal hypothyroidism can occur after tubal flushing with oil-based contrast (So et al., 2017), especially in women with subclinical hypothyroidism before HSG (Mekaru et al., 2008). With regards to neonatal safety, a Japanese cohort showed that infants born to mothers undergoing HSG with oil-based contrast before conceiving had a higher screening rate (2.4%) versus normal screening population (0.7%) (Satoh et al., 2015). Although there is limited evidence on these safety issues, they should be fully discussed during clinical consultations.

Research implications

Future trials should evaluate long-term fertility outcomes after tubal flushing. The effectiveness needs to be tested in future trials addressing different populations, including women with advanced age, anovulation or tubal factor infertility. Safety data on women and their offspring are also needed to address the short-term and long-term safety concerns.

The therapeutic effects of contrast media should also be tested in techniques other than HSG, including HyCoSy, HyFoSy, hydrolaparoscopy as well as laparoscopy, for instance, using tubal flushing with oil-based contrast after confirming tubal patency with HyCoSy, or performing pre-ovulatory tubal flushing without any imaging after confirmed tubal patency, followed by intrauterine insemination or timed intercourse as suggested in some studies (Edelstam et al., 2008; Maheux-Lacroix et al., 2016). This would be an interesting alternative treatment for IVF in women with unexplained infertility.

The mechanical effects of flushing on the fallopian tubes seem to be the most reasonable theory as the effects persist after several menstrual cycles post tubal flushing and such effects have been observed in both oil-based and water-based contrast in a recent cohort study (Dreyer et al., 2018). With regard to the difference between different contrast media, we hypothesize that for tubes with mucus plugs or debris, the higher viscosity of the oil-based contrast causes a better flushing effect, maybe due to a higher pressure during tubal flushing procedure. A recently study found that the treatment effect of oil-based contrast as compared

to water based contrast specifically occurred in women suffering from severe pain during tubal flushing (van Welie et al., 2018). The higher intrauterine pressure associated with the dislodgement of mucus plugs and debris might cause more pain. However, we acknowledge that such hypothesis is difficult to test in animal models or humans. Hypothesis on other mechanisms including the effects on endometrial receptivity should be further tested in future research.

CONCLUSIONS

In women with infertility undergoing fertility workup, tubal flushing using oil‐based contrast medium probably increases clinical pregnancy rates within 6 months and may increase subsequent live‐birth rates, compared to tubal flushing with water‐based contrast medium or no intervention. Available evidence on fertility outcomes beyond 6 months is inadequate to draw firm conclusions.

REFERENCES

Al-Fadhli R, Sylvestre C, Buckett W, Tulandi T. A randomized study of laparoscopic chromopertubation with lipiodol versus saline in infertile women. *Fertility & Sterility* 2006;85: 505-507.

Alper MM, Garner PR, Spence JE, Quarrington AM. Pregnancy rates after hysterosalpingography with oil- and water-soluble contrast media. *Obstetrics & Gynecology* 1986;68: 6-9.

Bateman BG, Nunley WC, Jr., Kitchin JD, 3rd. Intravasation during hysterosalpingography using oil-base contrast media. *Fertil Steril* 1980;34: 439-443.

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

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

de Boer AD, Vemer HM, Willemsen WN, Sanders FB. Oil or aqueous contrast media for hysterosalpingography: a prospective, randomized, clinical study. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 1988;28: 65-68.

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

Dreyer K, van Eekelen R, Tjon-Kon-Fat RI, van der Steeg JW, Steures P, Eijkemans M, van der Veen F, Hompes P, Mol B, van Geloven N. The therapeutic effect of hysterosalpingography in couples with unexplained subfertility: a post-hoc analysis of a prospective multi-centre cohort study. *Reprod Biomed Online* 2018.

Dreyer K, van Rijswijk J, Mijatovic V, Goddijn M, Verhoeve HR, van Rooij IAJ, Hoek A, Bourdrez P, Nap AW, Rijnsaardt-Lukassen HGM *et al.* Oil-Based or Water-Based Contrast for Hysterosalpingography in Infertile Women. *New England Journal of Medicine* 2017;376: 2043-2052.

Dusak A, Soydinc HE, Onder H, Ekinci F, Goruk NY, Hamidi C, Bilici A. Venous intravasation as a complication and potential pitfall during hysterosalpingography: re-emerging study with a novel classification. *J Clin Imaging Sci* 2013;3: 67.

Edelstam G, Sjosten A, Bjuresten K, Ek I, Wanggren K, Spira J. A new rapid and effective method for treatment of unexplained infertility. *Hum Reprod* 2008;23: 852-856.

Fang F, Bai Y, Zhang Y, Faramand A. Oil-based versus water-based contrast for hysterosalpingography in infertile women: a systematic review and meta-analysis of randomized controlled trials. *Fertil Steril* 2018;110: 153-160 e153.

Gillespie HW. The Therapeutic Aspect of Hysterosalpingography. *Br J Radiol* 1965;38: 301-302.

Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;12: 9.

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. *Res Synth Methods* 2012;3: 98-110.

Higgins JPT, Green, S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. The Cochrane Collaboration.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP *et al.* The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med* 2015;162: 777-784.

Izumi G, Koga K, Takamura M, Bo W, Nagai M, Miyashita M, Harada M, Hirata T, Hirota Y, Yoshino O *et al.* Oil-Soluble Contrast Medium (OSCM) for Hysterosalpingography Modulates Dendritic Cell and Regulatory T Cell Profiles in the Peritoneal Cavity: A Possible Mechanism by Which OSCM Enhances Fertility. *J Immunol* 2017;198: 4277-4284.

Johnson NP. Review of lipiodol treatment for infertility - an innovative treatment for endometriosis-related infertility? *Aust N Z J Obstet Gynaecol* 2014;54: 9-12.

Johnson NP, Farquhar CM, Hadden WE, Suckling J, Yu Y, Sadler L. The FLUSH trial--flushing with lipiodol for unexplained (and endometriosis-related) subfertility by hysterosalpingography: a randomized trial. *Human Reproduction* 2004;19: 2043-2051.

Johnson NP, Kwok R, Stewart AW, Saththianathan M, Hadden WE, Chamley LW. Lipiodol fertility enhancement: two-year follow-up of a randomized trial suggests a transient benefit in endometriosis, but a sustained benefit in unexplained infertility. *Hum Reprod* 2007;22: 2857-2862.

King EL, Herring JS. Sterility studies in private practice. *Am J Obstet Gynecol* 1949;58: 258-266.

Letterie GS, Rose GS. Pregnancy rates after the use of oil-based and water-based contrast media to evaluate tubal patency. *Southern Medical Journal* 1990;83: 1402-1403.

Lim CP, Hasafa Z, Bhattacharya S, Maheshwari A. Should a hysterosalpingogram be a first-line investigation to diagnose female tubal subfertility in the modern subfertility workup? *Hum Reprod* 2011;26: 967-971.

Lindborg L, Thorburn J, Bergh C, Strandell A. Influence of HyCoSy on spontaneous pregnancy: a randomized controlled trial. *Human Reproduction* 2009;24: 1075-1079.

Lindequist S, Justesen P, Larsen C, Rasmussen F. Diagnostic quality and complications of hysterosalpingography: oil- versus water-soluble contrast media--a randomized prospective study. *Radiology* 1991;179: 69-74.

Lindequist S, Rasmussen F, Larsen C. Use of iotrolan versus ethiodized poppy-seed oil in hysterosalpingography. *Radiology* 1994;191: 513-517.

Maheux-Lacroix S, Boutin A, Moore L, Bergeron ME, Bujold E, Laberge P, Lemyre M, Dodin S. Hysterosalpingosonography for diagnosing tubal occlusion in subfertile women: a systematic review with meta-analysis. *Hum Reprod* 2014;29: 953-963.

Maheux-Lacroix S, Dodin S, Moore L, Bujold E, Lefebvre J, Bergeron ME. Preovulatory uterine flushing with saline as a treatment for unexplained infertility: a randomised controlled trial protocol. *BMJ Open* 2016;6: e009897.

Mekaru K, Kamiyama S, Masamoto H, Sakumoto K, Aoki Y. Thyroid function after hysterosalpingography using an oil-soluble iodinated contrast medium. *Gynecol Endocrinol* 2008;24: 498-501.

Mikulska D, Kurzawa R, Rozewicka L. Morphology of in vitro sperm phagocytosis by rat peritoneal macrophages under influence of oily contrast medium (Lipiodol). *Acta Eur Fertil* 1994;25: 203-206.

Miyamoto Y, Tsujimoto T, Iwai K, Ishida K, Uchimoto R, Miyazawa T, Azuma H. Safety and pharmacokinetics of iotrolan in hysterosalpingography. Retention and irritability compared with Lipiodol. *Invest Radiol* 1995;30: 538-543.

Mohiyiddeen L, Hardiman A, Fitzgerald C, Hughes E, Mol BW, Johnson N, Watson A. Tubal flushing for subfertility. *Cochrane Database Syst Rev* 2015: CD003718.

National Institute for Health and Care Excellence. Fertility problems: assessment and treatment. *NICE guideline (CG156)* 2013.

Nugent D, Watson AJ, Killick SR, Balen AH, Rutherford AJ. A randomized controlled trial of tubal flushing with lipiodol for unexplained infertility. *Fertility & Sterility* 2002;77: 173-175.

Nunley WC, Jr., Bateman BG, Kitchin JD, 3rd, Pope TL, Jr. Intravasation during hysterosalpingography using oil-base contrast medium--a second look. *Obstet Gynecol* 1987;70: 309-312.

Ogata R, Nakamura G, Uchiumi Y, Yokoyama M, Watanabe Y, Nozaki M, Sano M, Nakano H. Therapeutic efficacy of hysterosalpingography (HSG) in infertility, a prospective, randomized, clinical study. *Japanese Journal of and Sterility* 1993;38: 91-94.

Practice Committee of the American Society for Reproductive Medicine. Diagnostic evaluation of the infertile female: a committee opinion. *Fertil Steril* 2015;103: e44-50.

Rasmussen F, Lindequist S, Larsen C, Justesen P. Therapeutic effect of hysterosalpingography: oil- versus water-soluble contrast media--a randomized prospective study. *Radiology* 1991;179: 75-78.

Salanti G, Ades AE, Ioannidis JP. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *J Clin Epidemiol* 2011;64: 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.

Satoh M, Aso K, Katagiri Y. Thyroid Dysfunction in Neonates Born to Mothers Who Have Undergone Hysterosalpingography Involving an Oil-Soluble Iodinated Contrast Medium. *Horm Res Paediatr* 2015;84: 370-375.

Sawatari Y, Horii T, Hoshiai H. Oily contrast medium as a therapeutic agent for infertility because of mild endometriosis. *Fertil Steril* 1993;59: 907-911.

Siegler AM. Dangers of hysterosalpingography. *Obstetrical and Gynecological Survey* 1967;22: 284-308.

So S, Yamaguchi W, Tajima H, Nakayama T, Tamura N, Kanayama N, Tawara F. The effect of oil and water-soluble contrast medium in hysterosalpingography on thyroid function. *Gynecol Endocrinol* 2017;33: 682-685.

Soules MR, Spadoni LR. Oil versus aqueous media for hysterosalpingography: a continuing debate based on many opinions and few facts. *Fertil Steril* 1982;38: 1-11.

Spring DB, Barkan HE, Pruyn SC. Potential therapeutic effects of contrast materials in hysterosalpingography: a prospective randomized clinical trial. Kaiser Permanente Infertility Work Group. *Radiology* 2000;214: 53-57.

Steiner AZ, Meyer WR, Clark RL, Hartmann KE. Oil-soluble contrast during hysterosalpingography in women with proven tubal patency. *Obstetrics & Gynecology* 2003;101: 109-113.

van Rijswijk J, Pham CT, Dreyer K, Verhoeve HR, Hoek A, de Bruin JP, Nap AW, Wang R, Lambalk CB, Hompes PGA *et al.* Oil-based or water-based contrast for hysterosalpingography in infertile women: a cost-effective analysis of a randomized controlled trial. *Fertil Steril* 2018;110: 754-760.

van Rijswijk J, van Welie N, Dreyer K, van Hooff MHA, de Bruin JP, Verhoeve HR, Mol F, Kleiman-Broeze KA, Traas MAF, Muijsers G *et al.* The FOAM study: is Hysterosalpingo foam sonography (HyFoSy) a cost-effective alternative for hysterosalpingography (HSG) in assessing tubal patency in subfertile women? Study protocol for a randomized controlled trial. *BMC Womens Health* 2018;18: 64.

van Welie N, Dreyer K, van Rijswijk J, Verhoeve HR, Goddijn M, Nap AW, Smeenk JM, Traas MA, Lambalk CB, Hompes P *et al.* Treatment effect of oil-based contrast at HSG is dependent on pain at HSG but not on volume of contrast. *Fertility and Sterility* 2018;110: e40.

Wang W, Zhou Q, Zhou X, Chen Z, Zhang H. Influence Factors on Contrast Agent Venous Intravasation During Transvaginal 4-Dimensional Hysterosalpingo-Contrast Sonography. *J Ultrasound Med* 2018.

Watson A, Vandekerckhove P, Lilford R, Vail A, Brosens I, Hughes E. A meta-analysis of the therapeutic role of oil soluble contrast media at hysterosalpingography: a surprising result? *Fertil Steril* 1994;61: 470-477.

Weir WC, Weir DR. Therapeutic value of salpingograms in infertility. *Fertil Steril* 1951;2: 514-522.

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

Yang KN, Yeh NG, Pan SB. The therapeutic effects of oil-soluble hysterosalpingography contrast medium following water-soluble hysterosalpingography contrast medium. *Chung Hua i Hsueh Tsa Chih - Chinese Medical Journal* 1989;44: 293-297.

Yun AJ, Lee PY. Enhanced fertility after diagnostic hysterosalpingography using oil-based contrast agents may be attributable to immunomodulation. *AJR Am J Roentgenol* 2004;183: 1725-1727.

Chapter 7

The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria?

This is a pre-copyedited, author-produced version of an article accepted for publication in *Human Reproduction* following peer review. The version of record Wang R, Mol BWJ. The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria? *Human Reproduction* 2017;32(2):261-264 is available online at:
https://academic.oup.com/humrep/article/32/2/261/2452298

Statement of Authorship

Title of Paper	The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria?
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	Wang R, Mol BWJ. The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria? Human Reproduction 2017;32(2):261-264.

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	Conceptualised and drafted the manuscript; and approved the final version.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	2/2/2017

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

iv. the candidate's stated contribution to the publication is accurate (as detailed above);

v. permission is granted for the candidate in include the publication in the thesis; and

vi. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Ben Willem J Mol		
Contribution to the Paper	Conceptualised and drafted the manuscript; and approved the final version.		
Signature		Date	2/2/2017

ABSTRACT

The Rotterdam criteria for polycystic ovary syndrome (PCOS) are used by a wide range of medical professionals and researchers. However, the development of these criteria was based on expert meetings, and not on evidence-based treatment guidance. Over the last decade, the Rotterdam criteria have been useful in guiding research, and a number of clinical studies on PCOS have been published consequently. We plead to revisit the Rotterdam criteria based on the available evidence in prognostic studies and randomised controlled trials (RCTs). In this opinion paper, we provide arguments of strengths and limitations of the Rotterdam criteria in guiding treatment selections and predicting prognoses in women with infertility. While the Rotterdam criteria have shown their advantages in predicting reproductive prognosis, the next step is to evaluate whether they can guide treatment choices in infertility as well as other health aspects. Based on available data in clinical studies, we would be able to answer whether the Rotterdam criteria are evidence-based criteria.

KEYWORDS

polycystic ovary syndrome, diagnostic criteria, Rotterdam criteria, anovulation, infertility

INTRODUCTION

Diagnosis (Greek διαγιγνωσκειν to distinguish or discern) is the ascription of a name to an illness and implies the distinction of illness or disease from health (Pearce, 2011). Diagnostic criteria for a certain disease or syndrome are composed of a collection of symptoms, signs, as well as biochemical, genetic, imaging and pathological findings. Diagnostic criteria help to classify a disease as present or absent and have several purposes. First, they aim to estimate the natural course of disease, which is important information for the person involved. Subsequently, and more important, when a diagnosis indicates that the natural course of a disease is expected to be suboptimal, diagnostic criteria can guide treatment decisions, which have to aim to modify this prognosis in a beneficial way, thus improving the outcomes for the patients.

These rules should also be applied to polycystic ovary syndrome (PCOS). In 1935, Stein and Leventhal (Stein and Leventhal, 1935) reported a series of seven women with polycystic ovaries and oligo/amenorrhea, later to be known as PCOS. The chief complaints of these women were oligo/amenorrhea with subfertility, hirsutism or lower abdominal pain. Out of the seven women in the report, five were infertile, three were obese and three had hirsutism. All the seven women in the report gained normal menstruation after wedge resection and two of them became pregnant (Stein and Leventhal, 1935). Thus, the initial diagnosis of polycystic ovaries were related to patients' outcomes.

Stein and Leventhal diagnosed polycystic ovaries with pneumoroentgenography and laparotomy (Stein and Leventhal, 1935). These diagnostic methods were abandoned with the advent of hormonal assays in the 1970s (Rebar et al., 1976; Yen, 1980; Yen et al., 1970) and the introduction of high-resolution real-time ultrasonography in the 1980s (Adams et al., 1985; Swanson et al., 1981).

In 1990, the first international conference of PCOS was held at National Institutes of Health (NIH). Based on a consensus questionnaire of the attendees, rather than clinical research data, the following diagnostic criteria were put forth: oligo-anovulation and

hyperandrogenism/hyperandrogenaemia in the absence of all other endocrinopathies (Zawadzki and Dunaif, 1992). In 2003, a group of experts expanded the diagnostic criteria to include polycystic ovaries seen at ultrasound as a third diagnostic marker and to allow for a diagnosis of PCOS if two of three criteria were met and the same endocrinopathies were excluded, known as the Rotterdam criteria (Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004a, b). Again, the Rotterdam criteria were on the basis of closed session consensus among attendees. These new definitions were then accepted by the European Society for Human Reproduction and Embryology (ESHRE) and the American Society for Reproductive Medicine (ASRM) (Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group, 2004a, b) and two further consensus workshops regarding infertility management and various women's health aspects of PCOS were published afterwards (Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group, 2012; Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2008). Although the Rotterdam criteria are controversial (Azziz, 2006, Franks, 2006) and the Androgen Excess Society (AES) has proposed a new set of diagnostic criteria in 2006 (Azziz et al., 2006), they are still the most widely adopted criteria by different guidelines (Legro et al., 2013; Teede et al., 2011; Vause et al., 2010) and are used by a wide range of obstetricians and gynaecologists as well as other specialists.

DO THE ROTTERDAM CRITERIA GUIDE TREATMENT SELECTIONS?

Oligo-anovulatory infertile women with PCOS are treated with lifestyle intervention, medical or surgical ovulation induction and eventually IVF (Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2008). Oligo-anovulatory women with or without hyperandrogenism/hyperandrogenaemia or polycystic ovarian morphology (PCOM) will not affect clinical decision-making in ovulation induction treatment choices according to current guidelines (Legro et al., 2013; National Collaborating Centre for Women's and Children's Health (UK), 2013; Teede et al., 2011; Thessaloniki ESHRE/ASRM-Sponsored

PCOS Consensus Workshop Group, 2008). Thus, the diagnosis of PCOS according to the Rotterdam criteria does not guide treatments of oligo-anovulatory infertile women with PCOS.

In contrast, the World Health Organization (WHO) classification of anovulation is more pragmatic to guide treatment selection. Based on a preliminary classification published in 1968 (Insler et al., 1968), the WHO classified anovulation into three groups (World Health Organisation, 1973) and this classification was also adapted by ESHRE (The ESHRE Capri Workshop Group, 1995): 1) Group I: Hypogonadotropic hypogonadal anovulation, 2) Group II: Normogonadotropic normoestrogenic anovulation, and 3) Group III: Hypergonadotropic hypoestrogenic anovulation.

The WHO classification guides treatment. Women with WHO I anovulation need pulsatile administration of gonadotrophin-releasing hormone or gonadotrophins with luteinizing hormone activity (National Collaborating Centre for Women's and Children's Health (UK), 2013). Women in WHO-II group are predominately women with PCOS, for which clomiphene citrate has been the long-standing first-line medical treatment (ESHRE Capri Workshop Group, 2012; Legro et al., 2013; National Collaborating Centre for Women's and Children's Health (UK), 2013; Teede et al., 2011), with metformin, a combination of the two (National Collaborating Centre for Women's and Children's Health (UK), 2013) or letrozole (Legro et al., 2013; NHMRC, 2015) as recent and probably superior alternatives. Gonadotropins and laparoscopic ovarian surgery are considered as second-line treatment (ESHRE Capri Workshop Group, 2012; National Collaborating Centre for Women's and Children's Health (UK), 2013; Teede et al., 2011). For women with WHO class III, it is not effective to apply any of the ovulation-inducing regimens (The ESHRE Capri Workshop Group, 1995) with oocyte donation is an established fertility treatment in women with premature ovarian insufficiency (ESHRE Guideline Group on POI et al., 2016).

DO THE ROTTERDAM CRITERIA PREDICT PROGNOSIS?

Predictive models for pregnancy outcomes in infertile women with WHO group II anovulation and PCOS have been reported (Imani et al., 2000; Imani et al., 1998; Imani et al., 2002; Kuang et al., 2015; Mulders et al., 2003; Rausch et al., 2009; van Wely et al., 2005). In WHO II anovulatory women treated with clomiphene, predicting factors for live birth include free androgen index (FAI), BMI, oligomenorrhoea and age (Imani et al., 2002), while predictors for ovulation include FAI, BMI, oligomenorrhoea, and mean ovarian volume(Imani et al., 2000; Imani et al., 1998). Serum insulin-like growth factor-I (IGF-I), testosterone, age (Mulders et al., 2003), oligomenorrhoea, duration of infertility and FAI (van Wely et al., 2005) can predict the chance of ongoing pregnancy in women treated with follicle stimulating hormone (FSH). More recently, the predicting value of patient characteristics such as age, BMI, hirsutism score, FAI, insulin, and duration of infertility on reproductive outcomes in women with PCOS had been further confirmed in the data of PPCOSI and PPCOSII trials (Kuang et al., 2015; Rausch et al., 2009).

The predicting factors in these models are consist of different baseline characteristics. Some of them, such as hirsutism score, FAI, ovarian volume and oligomenorrhoea are important components of the Rotterdam criteria. Therefore, the different phenotypes of PCOS and the different components of the Rotterdam criteria can help to predict the reproductive outcomes to some extent, although other predictors (e.g. insulin) are not included in the Rotterdam criteria. These predicting factors need to be confirmed in different populations in future studies.

Existing biochemical tests for PCOS have poor sensitivity and specificity (Iliodromiti et al., 2013). Anti-Müllerian hormone (AMH), a hormone produced by granulosa cells of ovarian follicles during the early stages (Broer et al., 2014), is a promising biomaker in PCOS. AMH may be a good substitute for PCOM (Dewailly et al., 2011; Eilertsen et al., 2012) and also a useful initial diagnostic test for PCOS (Iliodromiti et al., 2013). However, AMH is not included in the Rotterdam criteria.

INFERTILITY IS A TIP OF THE ICEBERG IN PCOS

In the previous arguments, we exclusively focused on anovulatory infertility. For couples with anovulatory infertility, the current WHO based classification can help clinicians guide treatment. We believe that in a similar way it should be evaluated if criteria that are used to diagnose PCOS can be used to guide treatment choices. The Rotterdam criteria have shown their advantages in predicting reproductive outcomes in women PCOS. The next step is to discover whether can guide clinical decision making on treatment selections. With the introduction of individual participant data (IPD) meta-analysis to this area, it could be possible to solve this problem based on the IPD in prevously published RCTs. Additionally, this should also be the subject of future RCTs, in which not only a treatment effect is assessed in a dichotomous way, but also is evaluated whether a treatment effect is dependent on baseline characteristics (Janes et al., 2011).

Apart from infertility, it is clear that women fulfilling the PCOS criteria are at increased risk of pregnancy complications, long-term cardiovascular disease and endometrial cancer (Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group, 2012). However, it is unclear to which component of the PCOS criteria these risks are specifically related. As such, while there is a need for more prognostic studies that indicate which of the PCOS criteria are predictive for these complications, as well as RCTs that evaluate which characteristics of PCOS women can be used as treatment selection markers. IPD meta-analyses on these aspects of PCOS are also necessary for future research, as these studies can guide researchers and clinicians to find target population of different interventions and therefore provide evidence of personalised PCOS care.

While current criteria for PCOS are based on expert meetings, we plead to revisit them based on the evidence in prognostic studies and RCTs. The Rotterdam criteria have been useful in guiding research and therefore a number of clinical studies have been published over the past decade, but they should be evaluated for both prognostic capacity and the capacity to guide treatment.

CONCLUSIONS

As with many other complex syndromes, PCOS does not have a single diagnostic marker to provide a gold standard for reference. The consensus-based diagnostic criteria for PCOS in the Rotterdam criteria have defined the disease and, as such, have been valuable both clinically and scientifically. Although the Rotterdam criteria were developed based on expert opinions, research evidence has shown their advantages in predicting reproductive outcomes. As a next step, they should be evaluated for their capacity to guide treatment. We then could revisit these criteria based on both prognostic characteristics and treatment selection markers. This should guide the future status of the Rotterdam criteria for PCOS as to whether they are evidence-based criteria.

REFERENCES

Adams J, Franks S, Polson DW, Mason HD, Abdulwahid N, Tucker M, Morris DV, Price J, Jacobs HS. Multifollicular ovaries: clinical and endocrine features and response to pulsatile gonadotropin releasing hormone. *Lancet* 1985;2: 1375-1379.

Amsterdam ESHRE/ASRM-Sponsored 3rd PCOS Consensus Workshop Group. Consensus on women's health aspects of polycystic ovary syndrome (PCOS). *Hum Reprod* 2012;27: 14-24.

Azziz R. Controversy in clinical endocrinology: diagnosis of polycystic ovarian syndrome: the Rotterdam criteria are premature. *J Clin Endocrinol Metab* 2006;91: 781-785.

Azziz R, Carmina E, Dewailly D, Diamanti-Kandarakis E, Escobar-Morreale HF, Futterweit W, Janssen OE, Legro RS, Norman RJ, Taylor AE *et al.* Positions statement: criteria for defining polycystic ovary syndrome as a predominantly hyperandrogenic syndrome: an Androgen Excess Society guideline. *J Clin Endocrinol Metab* 2006;91: 4237-4245.

Broer SL, Broekmans FJ, Laven JS, Fauser BC. Anti-Mullerian hormone: ovarian reserve testing and its potential clinical implications. *Hum Reprod Update* 2014;20: 688-701.

Dewailly D, Gronier H, Poncelet E, Robin G, Leroy M, Pigny P, Duhamel A, Catteau-Jonard S. Diagnosis of polycystic ovary syndrome (PCOS): revisiting the threshold values of follicle count on ultrasound and of the serum AMH level for the definition of polycystic ovaries. *Hum Reprod* 2011;26: 3123-3129.

Eilertsen TB, Vanky E, Carlsen SM. Anti-Mullerian hormone in the diagnosis of polycystic ovary syndrome: can morphologic description be replaced? *Hum Reprod* 2012;27: 2494-2502.

ESHRE Capri Workshop Group. Health and fertility in World Health Organization group 2 anovulatory women. *Hum Reprod Update* 2012;18: 586-599.

ESHRE Guideline Group on POI, Webber L, Davies M, Anderson R, Bartlett J, Braat D, Cartwright B, Cifkova R, de Muinck Keizer-Schrama S, Hogervorst E *et al.* ESHRE

Guideline: management of women with premature ovarian insufficiency. *Hum Reprod* 2016;31: 926-937.

Franks S. Controversy in clinical endocrinology: diagnosis of polycystic ovarian syndrome: in defense of the Rotterdam criteria. *J Clin Endocrinol Metab* 2006;91: 786-789.

Iliodromiti S, Kelsey TW, Anderson RA, Nelson SM. Can anti-Mullerian hormone predict the diagnosis of polycystic ovary syndrome? A systematic review and meta-analysis of extracted data. *J Clin Endocrinol Metab* 2013;98: 3332-3340.

Imani B, Eijkemans MJ, de Jong FH, Payne NN, Bouchard P, Giudice LC, Fauser BC. Free androgen index and leptin are the most prominent endocrine predictors of ovarian response during clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. *J Clin Endocrinol Metab* 2000;85: 676-682.

Imani B, Eijkemans MJ, te Velde ER, Habbema JD, Fauser BC. Predictors of patients remaining anovulatory during clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. *J Clin Endocrinol Metab* 1998;83: 2361-2365.

Imani B, Eijkemans MJ, te Velde ER, Habbema JD, Fauser BC. A nomogram to predict the probability of live birth after clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. *Fertil Steril* 2002;77: 91-97.

Insler V, Melmed H, Mashiah S, Monselise M, Lunenfeld B, Rabau E. Functional classification of patients selected for gonadotropic therapy. *Obstet Gynecol* 1968;32: 620-626.

Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. *Ann Intern Med* 2011;154: 253-259.

Kuang H, Jin S, Hansen KR, Diamond MP, Coutifaris C, Casson P, Christman G, Alvero R, Huang H, Bates GW *et al.* Identification and replication of prediction models for ovulation, pregnancy and live birth in infertile women with polycystic ovary syndrome. *Hum Reprod* 2015;30: 2222-2233.

Legro RS, Arslanian SA, Ehrmann DA, Hoeger KM, Murad MH, Pasquali R, Welt CK, Endocrine S. Diagnosis and treatment of polycystic ovary syndrome: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab* 2013;98: 4565-4592.

Mulders AG, Eijkemans MJ, Imani B, Fauser BC. Prediction of chances for success or complications in gonadotrophin ovulation induction in normogonadotrophic anovulatory infertility. *Reprod Biomed Online* 2003;7: 170-178.

National Collaborating Centre for Women's and Children's Health (UK). Fertility: Assessment and Treatment for People with Fertility Problems. *National Institute for Health and Clinical Excellence: Guidance* 2013.

NHMRC. Evidence-based guideline for the assessment and management of polycystic ovary syndrome (Updated August 2015 - Section 7.4 Aromatase inhibitors). *Australian Government NHMRC Website* 2015.

Pearce JM. Disease, diagnosis or syndrome? *Pract Neurol* 2011;11: 91-97.

Rausch ME, Legro RS, Barnhart HX, Schlaff WD, Carr BR, Diamond MP, Carson SA, Steinkampf MP, McGovern PG, Cataldo NA *et al.* Predictors of pregnancy in women with polycystic ovary syndrome. *J Clin Endocrinol Metab* 2009;94: 3458-3466.

Rebar R, Judd HL, Yen SS, Rakoff J, Vandenberg G, Naftolin F. Characterization of the inappropriate gonadotropin secretion in polycystic ovary syndrome. *J Clin Invest* 1976;57: 1320-1329.

Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome. *Fertil Steril* 2004a;81: 19-25.

Rotterdam ESHRE/ASRM-Sponsored PCOS consensus workshop group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004b;19: 41-47.

Stein I, Leventhal M. Amenorrhea associated with bilateral polycystic ovaries. *Am J Obstet Gynecol* 1935;29: 181-191.

Swanson M, Sauerbrei EE, Cooperberg PL. Medical implications of ultrasonically detected polycystic ovaries. *J Clin Ultrasound* 1981;9: 219-222.

Teede HJ, Misso ML, Deeks AA, Moran LJ, Stuckey BG, Wong JL, Norman RJ, Costello MF, Guideline Development Groups. Assessment and management of polycystic ovary syndrome: summary of an evidence-based guideline. *Med J Aust* 2011;195: S65-112.

The ESHRE Capri Workshop Group. Anovulatory infertility. The ESHRE Capri Workshop Group. *Hum Reprod* 1995;10: 1549-1553.

Thessaloniki ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Consensus on infertility treatment related to polycystic ovary syndrome. *Hum Reprod* 2008;23: 462-477.

van Wely M, Bayram N, van der Veen F, Bossuyt PM. Predicting ongoing pregnancy following ovulation induction with recombinant FSH in women with polycystic ovary syndrome. *Hum Reprod* 2005;20: 1827-1832.

Vause TD, Cheung AP, Sierra S, Claman P, Graham J, Guillemin JA, Lapensee L, Stewart S, Wong BC, Society of O *et al.* Ovulation induction in polycystic ovary syndrome. *J Obstet Gynaecol Can* 2010;32: 495-502.

World Health Organisation. Agents stimulating gonadal function in the human. Report of a WHO scientific group. *World Health Organ Tech Rep Ser* 1973;514: 1-30.

Yen SS. The polycystic ovary syndrome. *Clin Endocrinol (Oxf)* 1980;12: 177-207.

Yen SS, Vela P, Rankin J. Inappropriate secretion of follicle-stimulating hormone and luteinizing hormone in polycystic ovarian disease. *J Clin Endocrinol Metab* 1970;30: 435-442.

Zawadzki J, Dunaif A. Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In Dunaif A, Givens H, Haseltine F and Merriam G (eds) *Polycystic Ovary Syndrome*. 1992. Blackwell Scientific, Boston, MA, 377-384.

Chapter 8

First-line ovulation induction for polycystic ovary syndrome: an individual participant data meta-analysis

This is a pre-copyedited, author-produced version of an article accepted for publication in *Human Reproduction Update* following peer review. The version of record Wang R, Li W, Bordewijk EM, Legro RS, Zhang H, Wu X, Gao J, Morin-Papunen L, Homburg R, König TE, Moll E, Kar S, Huang W, Johnson NP, Amer SA, Vegetti W, Palomba S, Falbo A, Özmen Ü, Nazik H, Williams CD, Federica G, Lord J, Sahin Y, Bhattacharya S, Norman RJ, van Wely M, Mol BW, Reproductive Medicine Network, the International Ovulation Induction IPDMA Collaboration. First-line ovulation induction for polycystic ovary syndrome: an individual participant data meta-analysis. *Human Reproduction Update*. 2019 will be available online at: https://academic.oup.com/humupd

Statement of Authorship

Title of Paper	First-line ovulation induction for polycystic ovary syndrome – an individual participant data meta-analysis
Publication Status	☐ Published ☑ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted w ork w ritten in manuscript style
Publication Details	Wang R, Li W, Bordewijk EM, Legro RS, Zhang H, Wu X, Gao J, Morin-Papunen L, Homburg R, König TE, Moll E, Kar S, Huang W, Johnson NP, Amer SA, Vegetti W, Palomba S, Falbo A, Özmen Ü, Nazik H, Williams CD, Federica G, Lord J, Sahin Y, Bhattacharya S, Norman RJ, van Wely M, Mol BW, Reproductive Medicine Network, the International Ovulation Induction IPDMA Collaboration. First-line ovulation induction for polycystic ovary syndrome: an individual participant data meta-analysis. Human Reproduction Update. 2019. Accepted.

Principal Author

Name of Principal Author (Candidate)	Rui Wang		
Contribution to the Paper	Conceptualised and designed the study, collected the data, cleaned and analysed the data, drafted the first manuscript, interpreted the pooled data, and approved the final version.		
Overall percentage (%)	70%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	9/8/2019

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

vii. the candidate's stated contribution to the publication is accurate (as detailed above);

viii. permission is granted for the candidate in include the publication in the thesis; and

ix. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Wentao Li		
Contribution to the Paper	WL collected the data, cleaned and analysed the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

Name of Co-Author	Esmée M Bordewijk
Contribution to the Paper	EMB collected the data, cleaned and analysed the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.

Signature		Date	14/8/2019

Name of Co-Author	Richard S Legro		
Contribution to the Paper	RSL conceptualised and designed the study, provided and interpreted data of included trials, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Heping Zhang		
Contribution to the Paper	HZ provided and interpreted data of included trials, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Xiaoke Wu		
Contribution to the Paper	XW provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	14/8/2019

Name of Co-Author	Jingshu Gao		
Contribution to the Paper	JG provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	14/8/2019

Name of Co-Author	Laure Morin-Papunen		
Contribution to the Paper	L M-P provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Roy Homburg		
Contribution to the Paper	RH provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	14/8/2019

Name of Co-Author	Tamar E König		
Contribution to the Paper	TEK provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Etelka Moll		
Contribution to the Paper	EM provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	16/8/2019

Name of Co-Author	Sujata Kar		
Contribution to the Paper	SK provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	13/8/2019

Name of Co-Author	Wei Huang		
Contribution to the Paper	WH provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

Name of Co-Author	Neil P Johnson		
Contribution to the Paper	NPJ provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	13/8/2019

Name of Co-Author	Saad A Amer		
Contribution to the Paper	SAA provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	12/8/2019

Name of Co-Author	Walter Vegetti
Contribution to the Paper	WV provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.

Signature		Date	16/8/2019

Name of Co-Author	Stefano Palomba		
Contribution to the Paper	SP provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	16/8/2019

Name of Co-Author	Angela Falbo		
Contribution to the Paper	AF provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	16/8/2019

Name of Co-Author	Ülkü Özmen		
Contribution to the Paper	UO provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Hakan Nazik		
Contribution to the Paper	HN provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	14/8/2019

Name of Co-Author	Christopher D Williams		
Contribution to the Paper	CDW provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	15/8/2019

Name of Co-Author	Grasso Federica		
Contribution to the Paper	GF provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	14/8/2019

Name of Co-Author	Jonathan Lord		
Contribution to the Paper	JL provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	16/8/2019

Name of Co-Author	Yilmaz Sahin		
Contribution to the Paper	YS provided and interpreted data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	16/8/2019

Name of Co-Author	Siladitya Bhattacharya		
Contribution to the Paper	SB conceptualised and designed the study, collected the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

Name of Co-Author	Robert J Norman		
Contribution to the Paper	RJN conceptualised and designed the study, collected the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

Name of Co-Author	Madelon van Wely		
Contribution to the Paper	MvW conceptualised and designed the study, collected the data, cleaned and analysed the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

Name of Co-Author	Ben Willem Mol		
Contribution to the Paper	BWM conceptualised and designed the study, collected the data, cleaned and analysed the data, interpreted the pooled data, critically revised the manuscript for important intellectual content and approved the final version.		
Signature		Date	9/8/2019

ABSTRACT

Background: Polycystic ovary syndrome (PCOS) is the most frequent cause of anovulatory infertility. In women with PCOS, effective ovulation induction serves as an important first-line treatment for anovulatory infertility. Individual participant data (IPD) meta-analysis is considered as the gold standard for evidence synthesis which provides accurate assessments of outcomes from primary randomised controlled trials (RCTs) and allows additional analyses for time-to-event outcomes. It also facilitates treatment-covariate interaction analyses and therefore offers an opportunity for personalised medicine.

Objective and rationale: We aimed to evaluate the effectiveness of different ovulation induction agents, in particular letrozole alone and clomiphene citrate (CC) plus metformin, as compared to CC alone, as the first-line choice for ovulation induction in women with PCOS and infertility, and to explore interactions between treatment- and participant-level baseline characteristics.

Search methods: We searched electronic databases including MEDLINE, EMBASE and Cochrane Central Register of Controlled Trials up to 20th December 2018. We included RCTs comparing the following interventions with each other or placebo/ no treatment in women with PCOS and infertility: CC, metformin, CC plus metformin, letrozole, gonadotrophin and tamoxifen. We excluded studies on treatment-resistant women. The primary outcome was live birth. We contacted the investigators of eligible RCTs to share the IPD and performed IPD meta-analyses. We assessed the risk of bias by using the Cochrane risk of bias tool for RCTs.

Outcomes: IPD of 20 RCTs including 3962 women with PCOS were obtained. Six RCTs compared letrozole and CC in 1284 women. Compared with CC, letrozole improved live birth rates (3 RCTs, 1043 women, risk ratio [RR] 1.43, 95% confidence interval [CI] 1.17-1.75, moderate-certainty evidence) and clinical pregnancy rates (6 RCTs, 1284 women, RR 1.45, 95% CI 1.23-1.70, moderate-certainty evidence), and reduced time-to-pregnancy (6 RCTs, 1235 women, hazard ratio [HR] 1.72, 95%CI 1.38-2.15, moderate-certainty

evidence). Meta-analyses of effect modifications showed a positive interaction between baseline serum total testosterone levels and treatment effects on live birth (interaction RR 1.29, 95%CI 1.01-1.65).

Eight RCTs compared CC plus metformin to CC alone in 1039 women. Compared with CC alone, CC plus metformin might improve clinical pregnancy rates (8 RCTs, 1039 women, RR 1.18, 95% CI 1.00-1.39, low-certainty evidence) and might reduce time-to-pregnancy (7 RCTs, 898 women, HR 1.25, 95%CI 1.00-1.57, low-certainty evidence), but there was insufficient evidence of a difference on live birth rates (5 RCTs, 907 women, RR 1.08, 95% CI 0.87-1.35, low-certainty evidence). Meta-analyses of effect modifications showed a positive interaction between baseline insulin levels and treatment effects on live birth in the comparison between CC plus metformin and CC (interaction RR 1.03, 95% CI 1.01-1.06).

Wider implications: In women with PCOS, letrozole improves live birth and clinical pregnancy rates and reduces time-to-pregnancy compared to CC and therefore can be recommended as the preferred first-line treatment for women with PCOS and infertility. CC plus metformin may increase clinical pregnancy and may reduce time-to-pregnancy compared to CC alone, while there is insufficient evidence of a difference on live birth. Treatments effects of letrozole are influenced by baseline serum levels of total testosterone, while those of CC plus metformin are affected by baseline serum levels of insulin. These interactions between treatments and biomarkers on hyperandrogenaemia and insulin resistance provide further insights into a personalised approach for the management of anovulatory infertility related to PCOS.

KEY WORDS:

polycystic ovary syndrome, infertility, anovulation, ovulation induction, letrozole, clomiphene, metformin, individual participant data, meta-analysis.

INTRODUCTION

Polycystic ovary syndrome (PCOS) is the most common endocrine disorder of reproductive age women, and the prevalence among different geographic regions ranges from 5% to 21%, depending on the criteria used (Lizneva et al., 2016). PCOS is a heterogeneous syndrome comprising of at least two of the following clinical characteristics according to the Rotterdam diagnostic criteria: oligo-/ anovulation, clinical and/or biochemical hyperandrogenism, or polycystic ovaries morphology based on ultrasound assessment (Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group, 2004).

Anovulatory infertility is usually one of the key features that women with PCOS are confronted with. Simple and effective infertility treatments as the first-line choice are therefore important. Our previous network meta-analysis compared available first-line treatment options for women with PCOS with infertility and found that letrozole and combined clomiphene citrate (CC)-metformin were superior to other ovulation induction medications in terms of clinical pregnancy and that letrozole resulted in more live births than other interventions, including CC (Wang et al., 2017). These findings are in agreement with the evidence summarised in the International evidence based guideline for the assessment and management of PCOS (Teede et al., 2018).

As women with PCOS represent a heterogeneous population according to the diagnostic criteria, it is important to identify which individuals benefit most from a particular treatment so that clinicians can provide personalised care (Wang and Mol, 2017). However, primary RCTs are usually underpowered to detect subgroup effects (Riley et al., 2010). Subgroup analyses in meta-analyses of aggregate data are at risk of ecological bias due to the ignorance of within-study interactions, or are even impossible to perform due to heterogeneous reporting of subgroup data in the primary trials (Riley et al., 2010).

Moreover, time-to-pregnancy is also an important patient-centred outcome, but it has never been reported in previous meta-analyses on PCOS. This is likely due to the unavailability of the data in the publication as well as the methodological challenges on data extraction and

synthesis. In addition, the primary trials are not always of high quality in terms of analyses and reports (Eshre Capri Workshop Group, 2018), which can directly affect the data extraction, analysis and risk of bias assessment process in subsequent meta-analyses.

These deficiencies in aggregate data meta-analyses can potentially be overcome by using individual participant data (IPD). IPD meta-analysis has been described as the gold standard in evidence synthesis, by engaging investigators of the primary trials to provide the raw data of the primary trials (Broeze et al., 2010). Such strategy facilitates derivation of the information beyond the primary publication, standardisation of inclusion criteria, outcomes and analyses across trials, and investigations of subgroup effects and time-to-event outcomes. (Broeze et al., 2010; Riley et al., 2010).

We therefore performed an IPD meta-analysis to evaluate the effectiveness of different ovulation induction agents, in particular letrozole alone and CC plus metformin, as compared to CC alone, as the first-line choice for ovulation induction in women with PCOS and infertility, and to explore interactions between treatment- and participant-level baseline characteristics.

METHODS

Registration and literature search

This IPD meta-analysis was conducted based on a registered protocol (PROSPERO CRD42017059251) and reported according to the Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data (PRISMA-IPD) statement (Stewart et al., 2015).

We updated the searches in MEDLINE, EMBASE and Cochrane Central Register of Controlled Trials in September 2017, based on our previous search strategies for a network meta-analysis on treatment strategies for World Health Organization (WHO) II anovulation (Wang et al., 2017). In brief, the search terms included both index terms as well as free words on PCOS, anovulation and ovulation induction. After completing data requesting process, we further updated the search on 20th December 2018 to identify the latest studies. We also searched the WHO International Clinical Trials Registry Platform (WHO ICTRP) and U.S. National Institutes of Health (clinicaltrials.gov) and ISRCTN registry to identify ongoing trials. In addition, we reviewed the references lists of relevant papers and corresponded with trialists in PCOS to identify potential eligible trials that we might have missed.

Eligibility criteria

We included RCTs comparing the following interventions with each other or placebo/no treatment: clomiphene citrate (CC), metformin, CC and metformin combined, letrozole, gonadotrophins and tamoxifen in women with WHO II anovulation, including PCOS. We excluded trials reporting on treatment-resistant women, trials comparing different doses of the same intervention and quasi-RCTs. We did not apply language restrictions. For crossover trials, we only included the data in the first phase.

The primary outcome was live birth. The secondary outcomes were clinical pregnancy, ovulation, miscarriage, multiple pregnancy and time to pregnancy.

Study selection and data collection

Two members of the review team (from RW, WL and EMB) independently assessed the titles and abstracts to exclude irrelevant studies and subsequently reviewed the full-text articles to evaluate their eligibility. Disagreements were resolved by discussion with a third author (BWM, MvW or RJN).

We contacted investigators of eligible RCTs to share the de-identified IPD and established the International Ovulation Induction IPDMA Collaboration. We sent at least two more reminders when we did not receive responses.

We obtained de-identified IPD including baseline characteristics including age, body mass index (BMI), ethnicity, type of infertility (primary/secondary), treatment history (treatment-naïve or not), fasting glucose, fasting insulin, total testosterone, sex hormone binding globulin (SHBG), ovarian volume and the Ferriman-Gallwey score for hirsutism. We also obtained data on allocated treatments, number of ovulation induction cycles, ovulation and fertility outcomes including live birth, clinical pregnancy, miscarriage and multiple pregnancy.

We checked data for consistency by comparing the analyses from obtained IPD with the original publications. We discussed any inconsistencies or obvious errors with investigators of primary RCTs and solved discrepancies by consensus.

Risk of bias assessment

Two members of the review team independently evaluated the risk of bias in each included RCT, using the domain-based evaluation tool described in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green, 2011). We assessed the following domains as low risk of bias, unclear or high risk of bias: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete outcome data, selective reporting (reporting bias) and other sources of

bias. When the risk of bias for a domain was unclear, investigators of these RCTs were asked to provide additional information to resolve the uncertainty.

We assessed the overall certainty of the evidence across RCTs by using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach, including the risk of bias, consistency of effect, imprecision, indirectness and publication bias.

Data synthesis

We conducted all analyses based on an intention-to-treat principle using woman randomised per allocated group as the unit of all analyses. We performed two-stage random-effects IPD meta-analyses for letrozole versus CC alone and CC with metformin versus CC alone. For dichotomous outcomes, we calculated risk ratios (RRs) and 95% confidence intervals (CIs) and presented statistical heterogeneity by using I^2 statistic (Higgins and Green, 2011). For time-to-event outcomes, we used the number of treatment cycles as an approximate estimate for time and visualised the summary time-to-event in simple non-stratified Kaplan-Meier curves. We also estimated hazard ratios (HR) in Cox proportional hazards regression models for discrete time and pooled HRs and 95% CI, by using the generic inverse variance method (Fisher, 2015).

Subgroup effects were estimated for the primary outcome by treatment-covariate interaction terms within trials and subsequent meta-analyses of interactions, as interactions using within-trials information alone without considering between-trials interactions are recommended as the standard practice to avoid ecological bias (Fisher et al., 2017). We explored the treatment-covariate interactions of the following pre-specified baseline covariates: age, BMI, ethnicity, primary/secondary infertility, treatment history, hirsutism score, insulin resistance (serum glucose and insulin level), hyperandrogenaemia status (testosterone, SHBG, free androgen index) and ovarian volume. We also added the analysis of homeostatic model assessment for insulin resistance (HOMA-IR) as requested during the

peer review process. For dichotomous covariates with statistically significant interaction, we further performed stratified analyses to illustrate the treatment effects in different strata of the subgroups. Continuous variables were analysed as such without categorisation. For continuous covariates with statistically significant interaction, we further presented a weighted mean curve and pointwise confidence interval based on treatment-covariate interactions estimated in relevant studies. Due to the potential type I error, the results of subgroup analyses were all considered exploratory.

To evaluate the IPD availability bias, we performed a network meta-analysis of RCTs with IPD in a random-effects multivariate meta-analysis model (Riley et al., 2017; White, 2015) on live birth and clinical pregnancy, and then compared the results with a network meta-analysis of all eligible RCTs. If these results were consistent, we considered the included RCTs with IPD representative of all the eligible RCTs.

We performed a sensitivity analysis on studies with low risk of bias in allocation concealment as planned. As the majority of eligible studies focused only on treatment-naïve women with PCOS, these studies did not contribute to within-study interaction for treatment history and were not included in the treatment-covariate analysis. We performed a post-hoc sensitivity analysis by including only treatment-naïve women to demonstrate the robustness of the results.

We conducted all the analyses in Stata software version 15.1 (Stata Corp, College Station, TX, USA).

RESULTS

Characteristics of included studies

The final updated search yielded 709 non-duplicated studies (Figure 1). After screening the titles and abstracts, 636 irrelevant studies were excluded. Finally, a total of 62 studies (61 publications, 9356 women) fulfilled the inclusion criteria and were included. These studies were published in English (n=58), French (n=1) (Boudhraa et al., 2010), Italian (n=1) (Santonocito et al., 2009), Turkish (n=1) (Aygen et al., 2007) and Persian (n=1) (Lorzadeh et al., 2011).

IPD was not sought from eight studies (575 women), due to insufficient contact information (n=6; 359 women) (Beigi, 2006; Boudhraa et al., 2010; Cudmore and Tupper, 1966; El-Biely and Habba, 2001; Garcia et al., 1985; Johnson et al., 1966) or because the studies were identified after our data requesting timeline (n=2; 216 women) (Fatima et al., 2018; Topçu et al., 2017). For the remaining 54 studies (8781 women), the primary investigators were contacted to share IPD of the primary studies. IPD from 34 studies (4819 women) were not available, due to no response (n=23; 3258 women) (Abuelghar et al., 2013; Atay et al., 2006; Ayaz et al., 2013; Banerjee Ray et al., 2012; Basirat et al., 2012; Boostanfar et al., 2001; Chen et al., 2016; Dasari and Pranahita, 2009; Dehbashi et al., 2009; Hossein-Rashidi et al., 2016; Jahan, 2015; Karimzadeh et al., 2007; Karimzadeh and Javedani, 2010; Lopez et al., 2004; Lorzadeh et al., 2011; Maged et al., 2015; Robinson et al., 2003; Roy et al., 2012; Selim and Borg, 2012; Seyedoshohadaei et al., 2012; Sharief and Nafee, 2015; Sheikh-El-Arab Elsedeek and Elmaghraby, 2011; Zeinalzadeh et al., 2010), data loss (n=10; 1411 women) (Aygen et al., 2007; Badawy et al., 2009; Badawy and Gibreal, 2011; Fleming et al., 2002; Keikha and Shahraki, 2011; Khorram et al., 2006; Mobusher, 2014; Santonocito et al., 2009; Tang et al., 2006; Zain et al., 2009) or legal reasons (n=1; 150 women) (Moussa et al., 2016). The details of these studies are listed in Supplementary Table 1.

IPD were available for at least one outcome from 20 studies (3962 women, Table 1), including three from the US (Legro et al., 2007; Legro et al., 2014; Williams et al., 2009),

three from Italy (Leanza et al., 2014; Palomba et al., 2005; Vegetti et al., 1999), three from Turkey (Bayar et al., 2006; Nazik and Kumtepe, 2012; Sahin et al., 2004), two from the UK (Amer et al., 2017; Lord et al., 2006), two from China (Liu et al., 2017; Wu et al., 2017), two from India (Kar, 2012; Kar and Sanchita, 2015), two studies (in one publication) from New Zealand (Johnson et al., 2010), one from The Netherlands (Moll et al., 2006), one from Finland (Morin-Papunen et al., 2012) and one from multiple countries (The Netherlands, UK, Malta, Belgium, Argentina and Colombia) (Homburg et al., 2012). These RCTs were published in English between 1999 and 2017, with 11 (55%) published after 2010.

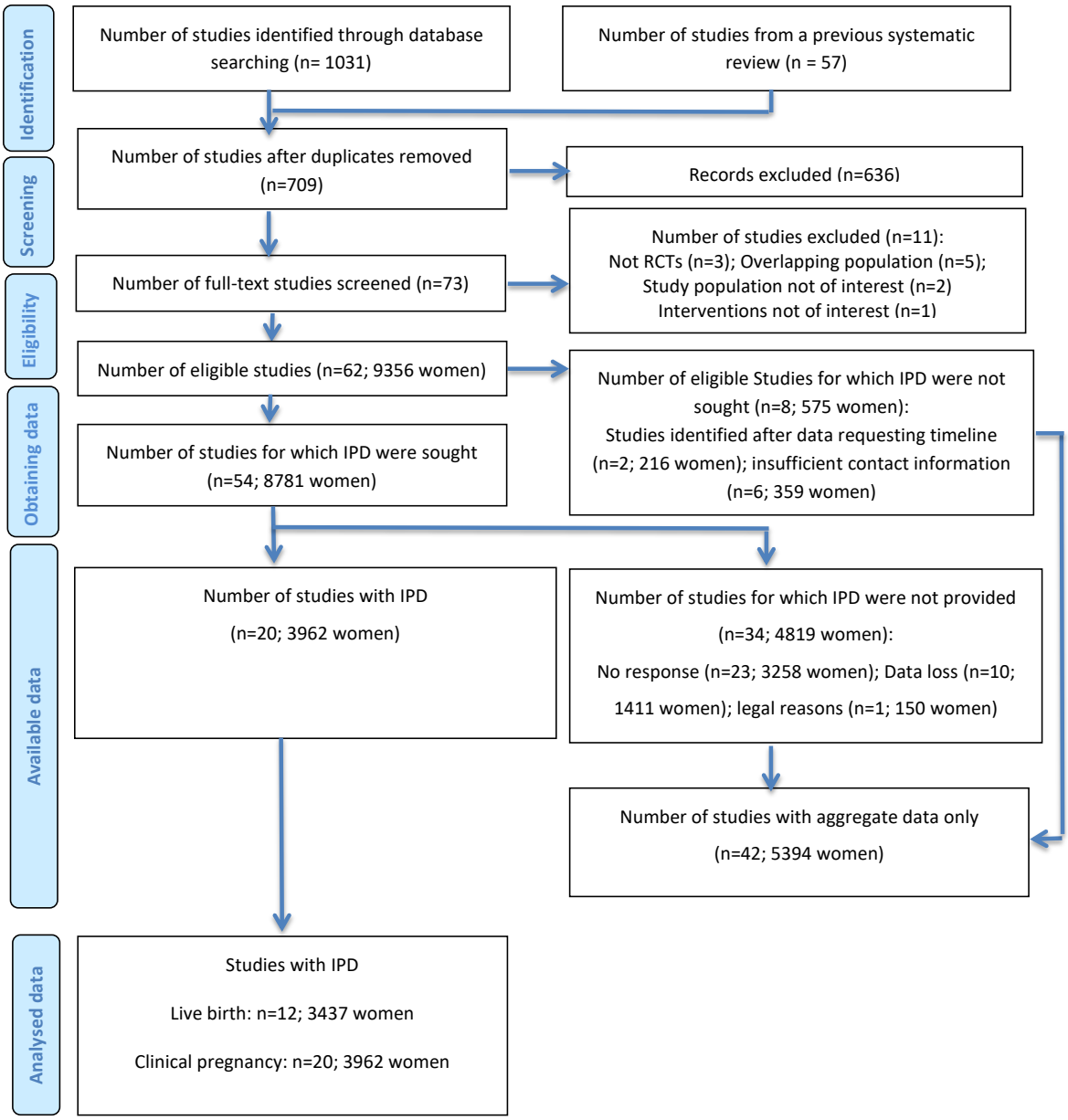


Figure 1. PRISMA-IPD flow diagram

Participants in all 20 RCTs were women with PCOS. In one RCT, participants were diagnosed with PCOS by fulfilling at least three of the following: PCO morphology, oligo/amenorrhoea, hirsutism, hyperandrogenaemia and elevated serum LH/FSH ratio (Sahin et al., 2004); while in the remaining 19 RCTs, the participants were women with PCOS based on the Rotterdam criteria (Bayar et al., 2006; Kar, 2012; Leanza et al., 2014; Liu et al., 2017; Nazik and Kumtepe, 2012) or different phenotypes, including Phenotype B (ovulatory dysfunction + androgen excess) (Amer et al., 2017; Homburg et al., 2012; Johnson et al., 2010; Kar and Sanchita, 2015; Legro et al., 2007; Legro et al., 2014; Lord et al., 2006; Morin-Papunen et al., 2012; Palomba et al., 2005; Williams et al., 2009; Wu et al., 2017) or Phenotype D (ovulatory dysfunction + PCO) (Moll et al., 2006; Vegetti et al., 1999).

For RCTs involving two stages of different interventions, including cross-over studies, we only included the data in the first stage. We included the IPD comparing letrozole versus CC before crossing over (Amer et al., 2017) and included the IPD comparing metformin versus placebo within the first three months before starting other ovulation induction agents (Morin-Papunen et al., 2012). In one RCT (Nazik and Kumtepe, 2012), switching between intervention and the control after the first cycle was allowed during the trial and the analysis in the primary publication was on a per-cycle basis; and therefore we only included the IPD of the first cycle.

In summary, four RCTs compared three interventions (CC plus metformin or CC alone versus metformin (Johnson et al., 2010; Kar and Sanchita, 2015; Legro et al., 2007) or CC with metformin or letrozole versus CC (Liu et al., 2017) and the remaining 16 compared two interventions. The most common comparisons were CC with metformin versus CC alone (8 RCTs) (Johnson et al., 2010; Kar and Sanchita, 2015; Leanza et al., 2014; Legro et al., 2007; Liu et al., 2017; Moll et al., 2006; Sahin et al., 2004; Williams et al., 2009) and letrozole versus CC alone (6 RCTs) (Amer et al., 2017; Bayar et al., 2006; Kar, 2012; Legro et al., 2014; Liu et al., 2017; Nazik and Kumtepe, 2012).

Table 1. Characteristics of included studies

Study	Comparisons	Sample Size	Age (mean)	BMI (mean)	Treatment-naïve (%)	Outcomes
Amer 2017	Letrozole vs CC	159	28.2±4.3	27.5±4.8	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Bayar 2006	Letrozole vs CC	80 (74)	31.4±4.0	NA	100%	clinical pregnancy, multiple pregnancy, time to pregnancy
Homburg 2012	FSH vs CC	302	29.5±3.9	25.4±5.6	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Johnson 2010A	Metformin vs placebo	65	29.6±4.2	37.8±3.5	69%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Johnson 2010B	CC+metformin vs CC vs Metformin	106	28.7±4.4	26.5±3.7	78%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Kar 2012	Letrozole vs CC	103	NA	25.9±3.4	100%	clinical pregnancy, time to pregnancy, ovulation
Kar 2015	CC+metformin vs CC vs Metformin	105 (81)	25.6±3.3	26.1±4.3	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Leanza 2014	CC+metformin vs CC	56	31.1±2.0	29.5±1.4	100%	clinical pregnancy, miscarriage, ovulation
Legro 2007	CC+metformin vs CC vs Metformin	626	28.1±4.0	35.2±8.7	45%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Legro 2014	Letrozole vs CC	750	28.9±4.3	35.1±9.3	45%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Liu 2017	CC+metformin vs letrozole vs CC	203	27.0±3.0	21.5±2.9	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Lord 2006	Metformin vs Placebo	44	29.1±4.9	34.8±7.0	unknown	Clinical pregnancy, ovulation
Moll 2006	CC+metformin vs CC	225	28.4±3.8	28.1±6.9	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Morin-Papunen 2012	Metformin vs Placebo	320	28.2±4.0	27.2±6.3	69%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Nazik 2012	Letrozole vs CC	64	26.8±5.6	25.1±4.3	100%	Clinical pregnancy, time to pregnancy, ovulation
Palomba 2005	CC vs metformin	100	26.2±4.4	26.7±2.3	100%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation
Sahin 2004	CC+metformin vs CC	21	25.1±3.3	28.2±3.7	100%	clinical pregnancy, time to pregnancy, ovulation
Vegetti 1999	Tamoxifen vs CC	95 (108)	30.9±3.1	22.7±4.2	100%	clinical pregnancy, time to pregnancy, ovulation
Williams 2009	CC+metformin vs CC	59 (55)	NA	NA	100%	clinical pregnancy, time to pregnancy, ovulation
Wu 2017	CC vs Placebo	500	27.9±3.3	24.5±4.2	70%	Live birth, clinical pregnancy, time to pregnancy, miscarriage, multiple pregnancy, ovulation

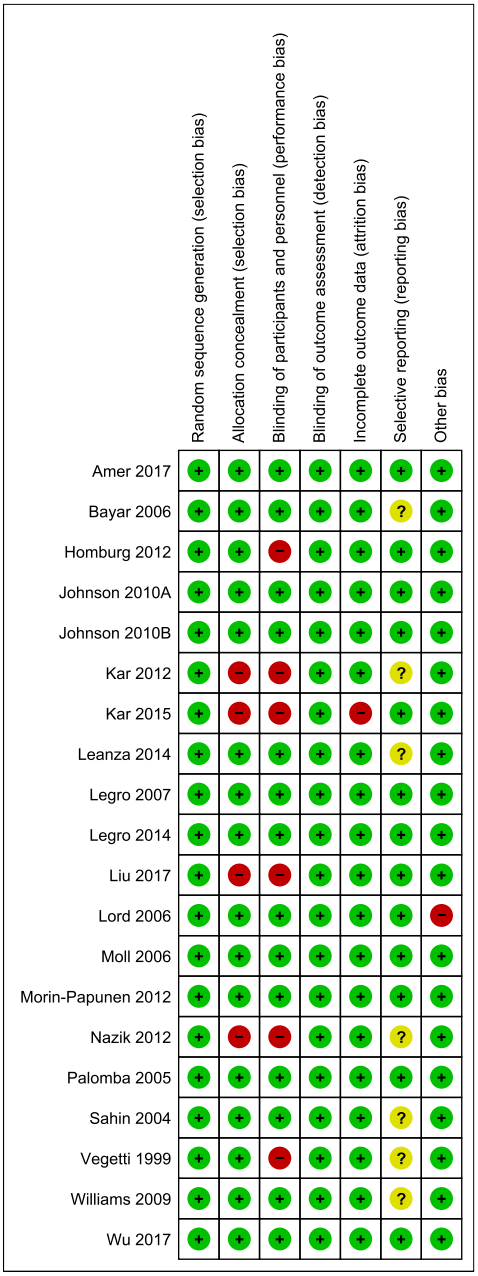


Figure 2. Risk of bias assessments of individual RCTs

Quality of evidence of individual studies

The details of risks of bias assessments within individual studies are presented in Figure 2. All RCTs (n=20) reported adequate methods of random sequence generation. Sixteen RCTs (80%) reported adequate methods of allocation concealment while the other four used an open allocation schedule without concealment (Kar, 2012; Kar and Sanchita, 2015; Liu et al., 2017; Nazik and Kumtepe, 2012). Fourteen RCTs (70%) blinded the participants and personnel during the trial while six RCTs applied an open label design (Homburg et al., 2012; Kar, 2012; Kar and Sanchita, 2015; Liu et al., 2017; Nazik and Kumtepe, 2012; Vegetti et al., 1999). Given that all outcomes of interest were objective outcomes, it is unlikely that

the non-blinded design will affect the outcome measurement and therefore detection bias was rated at low risk for all the included studies. One RCT (5%) had high risk of attrition bias, with 22% overall missing outcome data and 31% missing outcome data in the metformin group (Kar and Sanchita, 2015). One RCT (5%) was at another risk of bias due to allowing imbalanced co-intervention (CC) in both groups.

Meta-analyses of letrozole versus CC

Live birth

IPD were available in six RCTs comparing letrozole and CC, including 1284 women with PCOS. The forest plot of IPD Meta-analysis on live birth is presented in Figure 3a. Compared with CC, letrozole increased live birth rates (3 RCTs, 1043 women, RR 1.43, 95% CI 1.17-1.75, I^2=0, moderate certainty of evidence). Sensitivity analysis on studies with low risk of bias at allocation concealment and on treatment-naïve women were consistent with the main findings (2 RCTs, 909 women, RR 1.42, 95% CI 1.14-1.76, I^2=0; 3 RCTs, 627 women, RR 1.41, 95%CI 1.11-1.79, I^2=0) (Supplementary Table 2).

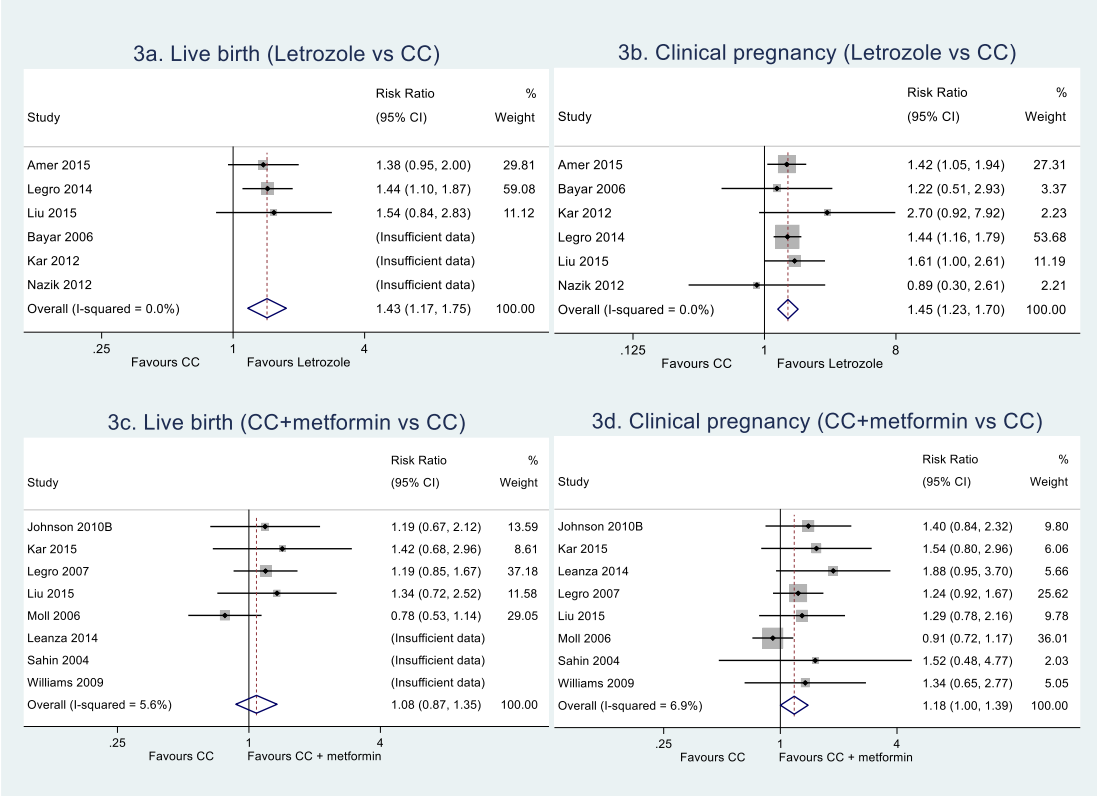


Figure 3. Meta-analyses of letrozole versus CC and CC plus metformin versus CC on live birth and clinical pregnancy

Secondary outcomes

Compared with CC alone, letrozole improved clinical pregnancy (6 RCTs, 1284 women, RR 1.45, 95%CI 1.23-1.70, I^2=0, moderate certainty of evidence, Figure 3b) and ovulation rates (5 RCTs, 1210 women, RR 1.13, 95%CI 1.07-1.20, I^2=0, moderate certainty of evidence, Table 2). There was insufficient evidence of a difference between letrozole and CC alone in terms of multiple pregnancy or miscarriage (Table 2).

The summary Kaplan-Meier curve for time to pregnancy is presented in Figure 4a. Subsequent pooled analysis of HRs showed that compared to CC, letrozole improved time-to-pregnancy (6 RCTs, 1235 women, HR 1.72, 95%CI 1.38-2.15, I^2=0, moderate certainty of evidence).

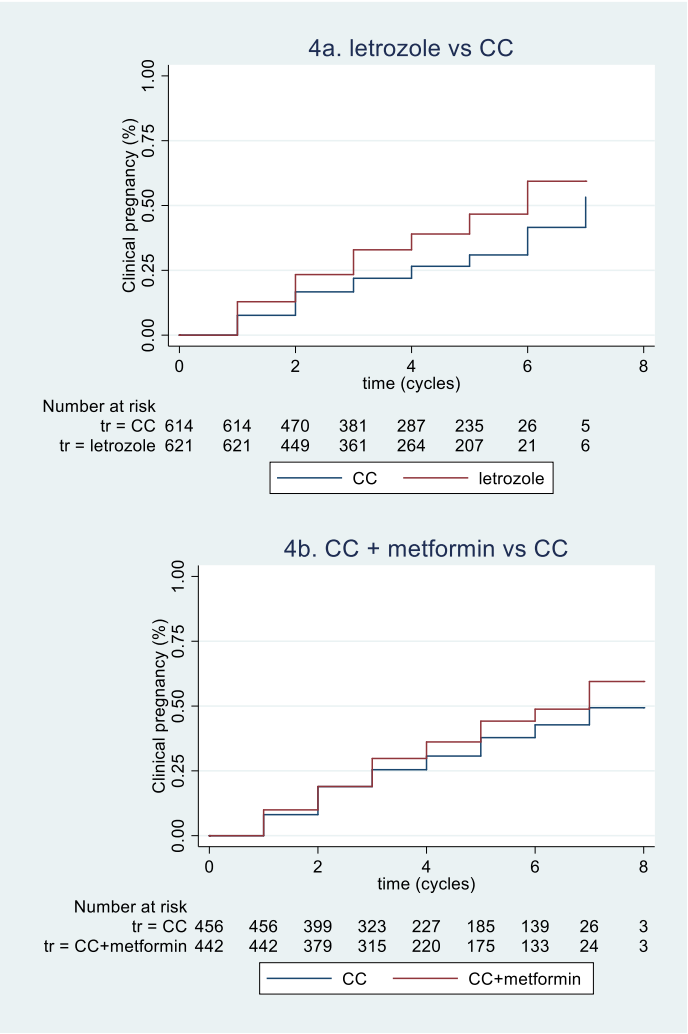


Figure 4. Summary Kaplan-Meier curves for time-to-event outcomes
Figure 4a-4b illustrate the non-stratified summary Kaplan-Meier curves for time-to-pregnancy in the comparisons of letrozole versus CC and CC plus metformin versus CC, respectively.
Participants with pregnancy before the first treatment cycles were not included in the 'Numbers at risk' table below and data were not stratified by trial in this Kaplan-Meier curve. The figures were intended to visualise time-to-event outcomes, but not to show statistical significance.

Treatment-covariate interactions

A meta-analyses of effect modifications showed a positive interaction between baseline serum total testosterone levels and treatment effects on live birth in the comparison between letrozole and CC (interaction RR 1.29, 95%CI 1.01-1.65, 3 RCTs, 1039 women, Figure 5a). This suggests that women with a higher baseline serum total testosterone level have a larger treatment effect of letrozole versus CC on live birth, compared to women with a lower baseline serum total testosterone level. Such an interaction was consistent across studies (I^2=0). To directly illustrate the association between baseline serum total testosterone level and relative treatment effects, this interaction is also presented in a weighted mean curve with 95% CI (Figure 5b). Meta-analyses did not find any other treatment-covariate interactions (Table 3).

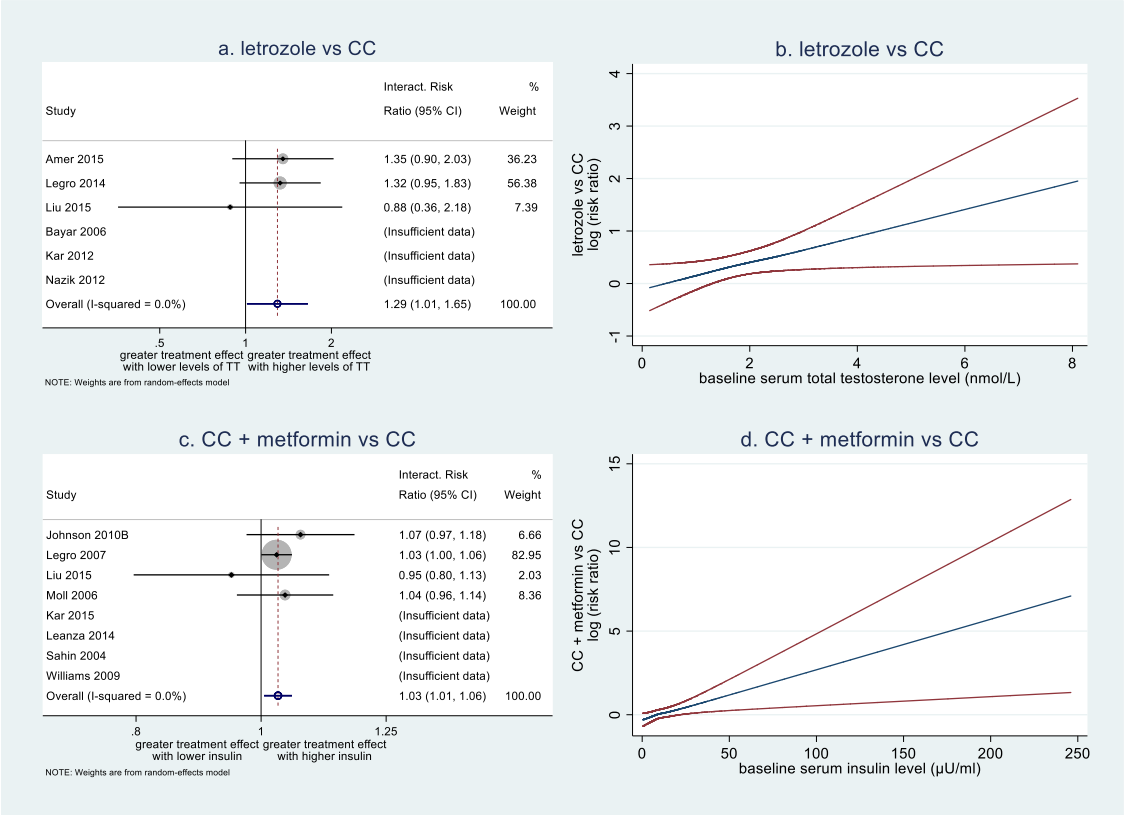


Figure 5. Forest plots and weighted mean curves for treatment-covariate interactions
5a. Forest plot of interactions between baseline serum total testosterone (TT) level and effect of letrozole versus CC on live birth. 5b. Weighted mean curve with pointwise 95% CI of interactions between baseline serum total testosterone level and relative effect of letrozole versus CC on live birth. 5c. Forest plot of interactions between baseline serum insulin level and effect of CC plus metformin versus CC on live birth. 5d. Weighted mean curve with pointwise 95% CIs of interactions between baseline serum insulin level and effect of CC plus metformin versus CC on live birth.
5a,c. Circles are used to depict the interaction effects within individual trials as well as the overall interaction effect. The sizes of the circles are in proportion to the inverse of the variance of the estimates. 5b,d. Blue line represents for the weighted mean effect of covariate on log risk ratios in the comparison between letrozole and CC. Red lines represent for pointwise 95% CI of interactions.

Table 2. Meta-analyses and GRADE assessments of all outcomes

Comparison	Outcome	Number of RCTs	Number of participants	Risk Ratio (RR)	95% confidence interval (CI)	I^2	Overall certainty of evidence (GRADE)
Letrozole vs CC	Live birth	3	1043	1.43	1.17-1.75	0	Moderate[1]
	Clinical pregnancy	6	1284	1.45	1.23-1.70	0	Moderate[1]
	Multiple pregnancy	2	909	1.45	0.17-12.45	50.9%	Very low[1,2,3]
	Miscarriage	3	1043	1.50	0.95-2.38	0	Low[1,3]
	Ovulation	5	1210	1.13	1.07-1.20	0	Moderate[1]
CC+metformin vs CC	Live birth	5	907	1.08	0.87-1.35	5.6%	Low[1,3]
	Clinical pregnancy	8	1039	1.18	1.00-1.39	6.9%	Low[1,3]
	Multiple pregnancy	4	771	0.76	0.24-2.42	0	Low[1,3]
	Miscarriage	6	963	1.33	0.79-2.26	0	Low[1,3]
	Ovulation	7	968	1.02	0.93-1.12	35.2%	Low[1,3]

1. Downgraded by one level due to concerns on risk of bias;
2. Downgraded by one level due to inconsistency;
3. Downgraded by one level due to imprecision.

Table 3. Meta-analyses of treatment-covariate interactions on live birth

Comparison	Baseline covariate	Number of RCTs	Number of participants	Interaction RR	Interaction 95% CI	Interaction I^2
Letrozole vs CC	Age	3	1043	0.98	0.93-1.05	24.9%
	BMI	3	1043	0.98	0.90-1.05	65.2%
	Ethnicity (non-Caucasian vs Caucasian)	2	909	1.42	0.80-2.45	0
	Treatment history (yes vs no)	1	750	1.07	0.63-1.82	/
	Type of infertility (secondary vs primary)	3	1043	0.83	0.43-1.60	52%
	Total testosterone (nmol/L)	3	1039	1.29	1.01-1.65	0
	SHBG (nmol/l)	2	907	1.00	0.99-1.02	69.7%
	Free androgen index	2	907	1.02	0.91-1.15	79.2%
	Fasting glucose (mmol/L)	3	1002	1.27	0.93-1.73	0
	Fasting insulin (µU/ml)	3	977	1.01	1.00-1.02	0
	HOMA-IR	3	975	1.04	0.98-1.09	0
	Ferriman–Gallwey score for hirsutism	2	884	1.03	0.99-1.06	0
	Ovarian volume (ml)	3	837	1.01	0.95-1.07	33.9%
CC + metformin vs CC	Age	5	895	1.06	0.98-1.15	43.7%
	BMI	5	885	1.02	0.98-1.07	25.1%
	Ethnicity (non-Caucasian vs Caucasian)	3	705	0.91	0.21-3.90	66.8%
	Treatment history (yes vs no)	1	418	0.90	0.46-1.78	/
	Type of infertility (secondary vs primary)	3	622	0.91	0.50-1.65	0
	Total testosterone (nmol/L)	4	824	1.02	0.95-1.08	0
	SHBG (nmol/l)	2	550	1.00	0.99-1.01	0
	Free androgen index	2	546	1.04	0.98-1.09	50.2%
	Fasting glucose (mmol/L)	4	812	1.01	0.74-1.37	0
	Fasting insulin (µU/ml)	4	741	1.03	1.01-1.06	0
	HOMA-IR	4	736	1.14	1.03-1.25	0
	Ferriman–Gallwey score for hirsutism	3	705	0.91	0.21-3.9	66.8%
	Ovarian volume (ml)	2	495	0.99	0.95-1.04	0

Meta-analyses of CC plus metformin versus CC

Live birth

IPD were available in eight RCTs comparing CC with metformin and CC alone, including 1039 women with PCOS. The forest plot of IPD Meta-analysis on live birth is presented in Figure 3c. Compared with CC alone, there was insufficient evidence of a difference between CC with metformin and CC alone on live birth (5 RCTs, 907 women, RR 1.08, 95%CI 0.87-1.35, I^2=5.6%, low certainty of evidence). Sensitivity analyses on studies with low risk of bias at allocation concealment and on treatment-naïve women showed very small treatment effects with wide CIs (3 RCTs, 714 women, RR 1.02, 95%CI 0.76-1.37, I^2=33.2%; 5 RCTs, 662 women, RR 1.06, 95%CI 0.83-1.34, I^2=3.9%) (Supplementary Table 2).

Secondary outcomes

Compared with CC alone, CC with metformin might improve clinical pregnancy (8 RCTs, 1039 women, RR 1.18, 95% CI 1.00-1.39, I^2=6.9%, low certainty of evidence, Figure 3b). There was insufficient evidence of a difference between CC with metformin and CC alone on ovulation, multiple pregnancy or miscarriage (Table 2).

The summary Kaplan-Meier curve is presented in Figure 4b. Pooled analysis of HRs showed that compared to CC alone, CC with metformin might improve time-to-pregnancy (7 RCTs, 898 women, HR 1.25, 95%CI 1.00-1.57, I^2=0, low certainty of evidence).

Treatment-covariate interactions

Meta-analyses of effect modifications showed a positive interaction between baseline insulin levels and treatment effects on live birth in the comparison between CC with metformin and CC alone (interaction RR 1.03, 95%CI 1.01-1.06, 4 RCTs, 741 women, Figure 5c). Such an interaction was consistent across studies (I^2=0). This suggests that women with a higher baseline serum insulin level have larger treatment effects of CC with metformin versus CC alone on live birth, compared to women with a lower baseline serum insulin level. Such an interaction was also presented in a weighted mean curve with 95%CI (Figure 5d). Additional meta-analysis of interactions for HOMA-IR was performed as requested during the peer

review process and it also showed a positive interaction between baseline HOMA-IR and treatment effects on live birth in the comparison between CC with metformin and CC alone (interaction RR 1.14, 95%CI 1.03-1.25, 4 RCTs, 736 women, I^2=0, Table 3). Meta-analyses did not find any other treatment-covariate interactions (Table 3).

IPD availability bias

With regards to IPD availability bias, network meta-analyses of 20 RCTs with IPD showed similar results to network meta-analyses of all eligible RCTs on both live birth and clinical pregnancy (Supplementary Table 3). Therefore, the participants in RCTs with IPD were representative of all the eligible participants with PCOS. The transitivity assumption of network meta-analyses was considered valid as the interventions of interest and placebo/no treatment were jointly randomisable.

DISCUSSION

Summary of evidence

This IPD meta-analysis showed that in women with PCOS, letrozole increased live birth rates compared to CC alone and the overall certainty of evidence was moderate. Such treatment benefits of letrozole compared to CC alone were more predominant in women with higher baseline serum levels of total testosterone. There was insufficient evidence of a difference between CC plus metformin and CC alone in live birth rates and the overall certainty of evidence was low, mainly due to risk of bias and imprecision. The potential benefit of CC in combination with metformin compared to CC alone were more pronounced in women with higher baseline serum insulin or HOMA-IR levels. We did not find other treatment-covariate interactions on live birth for other prespecified covariates including age, BMI, ethnicity, primary/secondary infertility, treatment history, Ferriman–Gallwey score for hirsutism, SHBG, free androgen index, fasting glucose levels or ovarian volume.

Strengths and limitations

Establishing the International Ovulation Induction IPDMA Collaboration facilitated a platform for key trialists in PCOS to collaborate and share the IPD of the primary trials. It provided us the opportunity to collect unpublished information of the primary trials including the details of randomisation and allocation concealment, treatment history, subgroup data and time-to-pregnancy. Such information allowed us to assess the quality of included trials precisely, to investigate treatment-covariate interactions and to take account of the time in the analyses. The findings of this IPD meta-analysis provide the best available up-to-date evidence.

Moreover, we applied a comprehensive search strategy without language restrictions and updated the search after completing data requesting in case we missed the most recent RCTs. Of the newly identified RCTs, one compared CC plus metformin vs CC in 128 women but did not report live birth (Fatima et al., 2018), while the other one compared tamoxifen vs CC

in 88 women (Topçu et al., 2017). Although we did not seek IPD from two RCTs identified after the data requesting deadline, adding IPD of these two studies is unlikely to change the main findings.

In addition, the investigation of subgroup effects includes within-study interaction only according to current statistical practice for IPD meta-analyses (Fisher et al., 2017) and therefore are free from ecological bias. For continuous covariates, without categorisation of the data, the statistical power was not compromised. Further illustration of interactions in weighted mean curve makes the interactions easier to interpret.

Nevertheless, this IPD meta-analysis has a few limitations. First, we were not able to access the IPD of all eligible studies. IPD were available for 32% (20/62) of the included trials, comprising 42% (3962/9356) of the eligible women with PCOS and the proportions of IPD availability was higher for studies reporting live birth (44% trials including 65% eligible women, Supplementary Table 3). This seems to be partly due to the long history of research on ovulation induction, with the first trial published in 1966. We were however able to access IPD of the highest-quality trials published within the last 15 years and we did not detect evidence of availability bias. Second, most of the planned subgroup analyses were based on two to three of the included studies and therefore may still be underpowered due to the unavailability of data on relevant covariates and/or live birth. Some primary trials only included a relatively homogeneous ethnicity group and therefore IPD in such trials could not contribute to the analysis of treatment-ethnicity interaction as no within-trial interaction was available. Third, as treatment-resistant women were excluded from this IPD meta-analysis, the findings can be applied in clinical practice on the choice of first-line treatment only. Last, we planned a one-stage IPD meta-analysis in the protocol but decided to use a two-stage approach before the final analysis. A two-stage approach allows graphical presentations for both overall treatment effects and treatment-covariate interactions, which is important for clinical interpretation, while it is not obvious how best to present graphically the results of a one-stage model (Fisher et al., 2017). In addition, the two-stage approach automatically

avoids ecological bias by accounting for within-trial interactions only (Fisher et al., 2017). Given the relatively large number of participants, low heterogeneity and overall good to moderate quality of included studies, we would expect both approaches to give very similar results.

Interpretations and clinical implications

The overall effects of letrozole and CC plus metformin vs CC on live birth and clinical pregnancy in this IPD meta-analysis were in agreement with existing systematic reviews (Franik et al., 2018; Morley et al., 2017; Wang et al., 2017) as well as the most recent the international evidence-based guideline recommendations (Teede et al., 2018). Based on the findings of this IPD meta-analysis, letrozole can be recommended as the first-line ovulation induction medication in women with PCOS and infertility, provided off-label use is allowed and women are fully informed. Compared to CC alone, CC plus metformin may increase clinical pregnancy rates but the evidence on live birth was insufficient. Sensitivity analysis showed that the treatment effects on live birth seemed very small. The discrepancies between clinical pregnancy and live birth were likely due to the bias arising from low quality of studies which did not report live birth. Further evidence is needed to address this question.

Subgroup analyses showed that women with higher baseline serum levels of total testosterone may benefit more from letrozole compared to CC and women with higher baseline serum levels of insulin may benefit more from CC plus metformin compared to CC alone. Such positive interactions were consistent across trials and supported from a biological perspective. Letrozole has been introduced as an ovulation induction agent since 2001 and it inhibits aromatase, therefore increasing gonadotropin secretion by release of the hypothalamic/pituitary axis from estrogenic negative feedback and resulting in stimulation of ovarian follicle development (Mitwally and Casper, 2001). According to the recent “two triangles hypothesis” for folliculogenesis in PCOS, pre-antral follicle growth is excessive due to intrinsic androgen excess that renders granulosa cells hypersensitive to FSH, with

consequently excessive AMH expression (Dewailly et al., 2016) Therefore, hyperandrogenaemia may improve the response to letrozole by enhancing the sensitivity of FSH receptors. However, such an interaction was not observed in other biomarkers of hyperandrogenaemia or hirsutism. This is likely due to the fact that the severity of hirsutism does not correlate well with the magnitude of androgen excess, as hirsutism is an expression of hyperandrogenism on hair follicles mediated through different pathways from those affecting the ovaries and follicles (Escobar-Morreale et al., 2012). Metformin is an insulin sensitising agent that decreases gluconeogenesis and lipogenesis and enhances peripheral glucose uptake and therefore increases insulin sensitivity (Naderpoor et al., 2015). The addition of metformin may further improve insulin resistance in women with higher fasting insulin or HOMA-IR levels and therefore improve pregnancy outcomes. We acknowledge that insulin levels are affected by many factors, ranging from physical activity and pre-test duration of fasting to sample handling and assay variability (Cassar et al., 2016). Therefore the international evidence-based guideline does not recommend clinical measurement of insulin resistance at present due to the lack of accuracy (Teede et al., 2018). In addition, SHBG has been proposed as a measure of insulin resistance (Cassar et al., 2016), but the findings in our IPD meta-analysis did not support treatment-by-SHBG interactions. Our work provides preliminary evidence that there may be a role for assessing insulin resistance in PCOS and infertility and supports the need to assess insulin resistance in infertility studies. We did not find ethnicity differences on treatment effects. This could be partly due to self-reported ethnicity without objective or DNA validation in all trials. We also did not find other treatment-covariate interactions on live birth for other prespecified covariates including age, BMI, primary/secondary infertility, treatment history, Ferriman–Gallwey score for hirsutism, SHBG, free androgen index, fasting glucose levels or ovarian volume. Although analyses of subgroup effects were prespecified in the protocol, these results should still be considered exploratory due to multiplicity.

Time is an important measurement for infertility outcomes, especially in the assessment of the effectiveness of multi-cycle treatments. However, time-to-event outcomes have seldomly been reported in meta-analyses of infertility trials as fertility outcomes are usually considered as dichotomous outcomes and Kaplan-Meier curves are rarely presented. Our IPD meta-analysis used number of cycles as a measure of time and evaluated time-to-pregnancy by estimating HRs and presenting summary Kaplan-Meier curves. Time-to-event analysis takes time and censored participants into account and provides more accurate estimates of treatment effect. Our analyses on time-to-pregnancy were inconsistent with those of clinical pregnancy.

Research implications

IPD meta-analyses are useful to inform the design, conduct, analysis, and interpretation of trials (Tierney et al., 2015). Given the consistent treatment benefits of letrozole across different fertility outcomes, future trials investigating new interventions for PCOS should choose letrozole as the reference arm. New trials are encouraged to incorporate treatment selection markers in their design to guide treatment decision (Janes et al., 2011), and the impact of these, including age, BMI and other biomarkers, need to be confirmed in future trials. More specifically, biomarkers for hyperandrogenaemia and insulin resistance could be applied in trials that evaluate metformin. Due to the limited accuracy for measuring existing insulin resistance biomarkers, optimal methods to assess insulin resistance in future trials should also be considered.

Developing and implementing a core outcome set for infertility (Duffy et al., 2018) and PCOS should be recommended to ensure outcomes are reported and collected consistently across future trials on infertility and PCOS to reduce research waste.

CONCLUSIONS

Our IPD meta-analysis shows that in women with PCOS, letrozole improves live birth and clinical pregnancy rates and reduces time-to-pregnancy compared to CC alone. CC plus metformin may improve clinical pregnancy rates and may reduce time-to-pregnancy compared to CC alone, but there is insufficient evidence of a difference on live birth.

Treatments effects of letrozole are influenced by baseline serum levels of total testosterone while those of CC plus metformin are affected by baseline serum levels of insulin. These interactions between treatments and biomarkers on hyperandrogenaemia and insulin resistance provide further insights into a personalised approach towards the clinical management of anovulatory infertility related to PCOS and therefore should be confirmed in future studies.

REFERENCES

Abuelghar WM, Elkady OS, Khamees AA. Clomiphene citrate alone, in combination with metformin or in combination with pioglitazone as first line therapy in induction of ovulation in infertile women with polycystic ovary syndrome, a randomized controlled trial. *Middle East Fertility Society Journal* 2013;18: 135-141.

Amer SA, Smith J, Mahran A, Fox P, Fakis A. Double-blind randomized controlled trial of letrozole versus clomiphene citrate in subfertile women with polycystic ovarian syndrome. *Human Reproduction* 2017;32: 1631-1638.

Atay V, Cam C, Muhcu M, Cam M, Karateke A. Comparison of letrozole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. *Journal of international medical research* 2006;34: 73-76.

Ayaz A, Alwan Y, Farooq MU. Metformin-clomiphene citrate vs. clomiphene citrate alone: Polycystic ovarian syndrome. *Journal of Human Reproductive Sciences* 2013;6: 15-18.

Aygen EM, Güzel Z, Özgün T, Atakul T, Şahin Y. The use of letrozole for ovulation induction in infertile women with polycystic ovarian syndrome [in Turkish]. *Erciyes Tip Dergisi* 2007;29: 195-200.

Badawy A, Abdel Aal I, Abulatta M. Clomiphene citrate or letrozole for ovulation induction in women with polycystic ovarian syndrome: a prospective randomized trial. *Fertility & Sterility* 2009;92: 849-852.

Badawy A, Gibreal A. Clomiphene citrate versus tamoxifen for ovulation induction in women with PCOS: a prospective randomized trial. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2011;159: 151-154.

Banerjee Ray P, Ray A, Chakraborti PS. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction in Indian women with polycystic ovarian syndrome. *Archives of Gynecology & Obstetrics* 2012;285: 873-877.

Basirat Z, Kashifard M, Amiri MG. Enhanced ovarian folliclular development by Metformin does not correlate with pregnancy rate: A randomized trial. *International Journal of Fertility and Sterility* 2012;6: 31-36.

Bayar U, Basaran M, Kiran S, Coskun A, Gezer S. Use of an aromatase inhibitor in patients with polycystic ovary syndrome: a prospective randomized trial. *Fertility & Sterility* 2006;86: 1447-1451.

Beigi A. Randomized trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in polycystic ovary syndrome. *Human reproduction (Oxford, England)* 2006;21: i129.

Boostanfar R, Jain JK, Mishell DR, Jr., Paulson RJ. A prospective randomized trial comparing clomiphene citrate with tamoxifen citrate for ovulation induction. *Fertility & Sterility* 2001;75: 1024-1026.

Boudhraa K, Jellouli MA, Amri M, Farhat M, Torkhani F, Gara MF. Indication of metformin in the management of hormonal dysfunction secondary to polycystic ovarian syndrome: prospective comparative study of 63 cases[in French]. *Tunisie medicale* 2010;88: 335-340.

Broeze KA, Opmeer BC, van der Veen F, Bossuyt PM, Bhattacharya S, Mol BW. Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. *Hum Reprod Update* 2010;16: 561-567.

Cassar S, Misso ML, Hopkins WG, Shaw CS, Teede HJ, Stepto NK. Insulin resistance in polycystic ovary syndrome: a systematic review and meta-analysis of euglycaemic-hyperinsulinaemic clamp studies. *Hum Reprod* 2016;31: 2619-2631.

Chen Z, Zhang M, Qiao Y, Yang J. Effects of letrozole in combination with low-dose intramuscular injection of human menopausal gonadotropin on ovulation and pregnancy of 156 patients with polycystic ovary syndrome. *Pakistan Journal of Medical Sciences* 2016;32: 1434-1438.

Cudmore DW, Tupper WR. Induction of ovulation with clomiphene citrate. A double-blind study. *Fertil Steril* 1966;17: 363-373.

Dasari P, Pranahita GK. The efficacy of metformin and clomiphene citrate combination compared with clomiphene citrate alone for ovulation induction in infertile patients with PCOS. *Journal of Human Reproductive Sciences* 2009;2: 18-22.

Dehbashi S, Dehbashi S, Kazerooni T, Robati M, Alborzi S, Parsanezhad ME, Shadman A. Comparison of the effects of letrozole and clomiphene citrate on ovulation and pregnancy rate in patients with polycystic ovary syndrome. *Iranian Journal of Medical Sciences* 2009;34: 23-28.

Dewailly D, Robin G, Peigne M, Decanter C, Pigny P, Catteau-Jonard S. Interactions between androgens, FSH, anti-Mullerian hormone and estradiol during folliculogenesis in the human normal and polycystic ovary. *Hum Reprod Update* 2016;22: 709-724.

Duffy JMN, Bhattacharya S, Curtis C, Evers JLH, Farquharson RG, Franik S, Khalaf Y, Legro RS, Lensen S, Mol BW *et al.* A protocol developing, disseminating and implementing a core outcome set for infertility. *Hum Reprod Open* 2018;2018: hoy007.

El-Biely MM, Habba M. The use of metformin to augment the induction of ovulation in obese infertile patients with polycystic ovary syndrome. *Middle East Fertility Society Journal* 2001;6: 43-49.

Escobar-Morreale HF, Carmina E, Dewailly D, Gambineri A, Kelestimur F, Moghetti P, Pugeat M, Qiao J, Wijeyaratne CN, Witchel SF *et al.* Epidemiology, diagnosis and management of hirsutism: a consensus statement by the Androgen Excess and Polycystic Ovary Syndrome Society. *Hum Reprod Update* 2012;18: 146-170.

Eshre Capri Workshop Group. Protect us from poor-quality medical research. *Hum Reprod* 2018;33: 770-776.

Fatima A, Khan SA, Saifuddin Z, Aslam R. Comparison of efficacy of clomiphene citrate alone and with metformin for treatment of infertility in polycystic ovarian syndrome. *Rawal Medical Journal* 2018;43: 285-288.

Fisher DJ. Two-stage individual participant data meta-analysis and generalized forest plots. *Stata Journal* 2015;15: 369-396.

Fisher DJ, Carpenter JR, Morris TP, Freeman SC, Tierney JF. Meta-analytical methods to identify who benefits most from treatments: daft, deluded, or deft approach? *BMJ* 2017;356: j573.

Fleming R, Hopkinson ZE, Wallace AM, Greer IA, Sattar N. Ovarian function and metabolic factors in women with oligomenorrhea treated with metformin in a randomized double blind placebo-controlled trial. *Journal of Clinical Endocrinology & Metabolism* 2002;87: 569-574.

Franik S, Eltrop SM, Kremer JA, Kiesel L, Farquhar C. Aromatase inhibitors (letrozole) for subfertile women with polycystic ovary syndrome. *Cochrane Database Syst Rev* 2018;5: CD010287.

Garcia CR, Freeman EW, Rickels K, Wu C, Scholl G, Galle PC, Boxer AS. Behavioral and emotional factors and treatment responses in a study of anovulatory infertile women. *Fertility & Sterility* 1985;44: 478-483.

Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. 2011. The Cochrane Collaboration.

Homburg R, Hendriks ML, Konig TE, Anderson RA, Balen AH, Brincat M, Child T, Davies M, D'Hooghe T, Martinez A *et al.* Clomifene citrate or low-dose FSH for the first-line treatment of infertile women with anovulation associated with polycystic ovary syndrome: a prospective randomized multinational study. *Human Reproduction* 2012;27: 468-473.

Hossein-Rashidi B, Khandzad B, Shahrokh-Tehraninejad E, Bagheri M, Gorginzadeh M. Recombinant FSH compared to clomiphene citrate as the first-line in ovulation induction in polycystic ovary syndrome using newly designed pens: A randomized controlled trial. *Journal of Family and Reproductive Health* 2016;10: 42-48.

Jahan S. Comparative study of efficacy among metformin, clomiphene citrate and aromatase inhibitor (letrozole) as the first-line medication for ovulation induction, achievement of pregnancy and live birth in Asian women with polycystic ovarian syndrome: A prospective trial. *International Journal of Gynecology and Obstetrics* 2015;131: E503.

Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. *Ann Intern Med* 2011;154: 253-259.

Johnson JE, Jr., Cohen MR, Goldfarb AF, Rakoff AE, Kistner RW, Plotz EJ, Vorys N. The efficacy of clomiphene citrate for induction of ovulation. A controlled study. *Int J Fertil* 1966;11: 265-270.

Johnson NP, Stewart AW, Falkiner J, Farquhar CM, Milsom S, Singh VP, Okonkwo QL, Buckingham KL, React-Nz am-cftg. PCOSMIC: a multi-centre randomized trial in women with PolyCystic Ovary Syndrome evaluating Metformin for Infertility with Clomiphene. *Human Reproduction* 2010;25: 1675-1683.

Kar S. Clomiphene citrate or letrozole as first-line ovulation induction drug in infertile PCOS women: A prospective randomized trial. *Journal of Human Reproductive Sciences* 2012;5: 262-265.

Kar S, Sanchita S. Clomiphene citrate, metformin or a combination of both as the first line ovulation induction drug for Asian Indian women with polycystic ovarian syndrome: A randomized controlled trial. *Journal of Human Reproductive Sciences* 2015;8: 197-201.

Karimzadeh MA, Eftekhar M, Taheripanah R, Tayebi N, Sakhavat L, Zare F. The effect of administration of metformin on lipid profile changes and insulin resistance in patients with polycystic ovary syndrome. *Middle East Fertility Society Journal* 2007;12: 174-178.

Karimzadeh MA, Javedani M. An assessment of lifestyle modification versus medical treatment with clomiphene citrate, metformin, and clomiphene citrate-metformin in patients with polycystic ovary syndrome. *Fertility & Sterility* 2010;94: 216-220.

Keikha F, Shahraki MB. Induction ovulation in polycystic ovary patient with clomiphene citrate and letrozole. *Iranian Journal of Reproductive Medicine* 2011;9: 46.

Khorram O, Helliwell JP, Katz S, Bonpane CM, Jaramillo L. Two weeks of metformin improves clomiphene citrate-induced ovulation and metabolic profiles in women with polycystic ovary syndrome. *Fertility & Sterility* 2006;85: 1448-1451.

Leanza V, Coco L, Grasso F, Leanza G, Zarbo G, Palumbo M. Ovulation induction with clomiphene citrate and metformin in women with polycystic ovary syndrome. *Minerva ginecologica* 2014;66: 299-301.

Legro RS, Barnhart HX, Schlaff WD, Carr BR, Diamond MP, Carson SA, Steinkampf MP, Coutifaris C, McGovern PG, Cataldo NA *et al.* Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. *New England journal of medicine* 2007;356: 551-566.

Legro RS, Brzyski RG, Diamond MP, Coutifaris C, Schlaff WD, Casson P, Christman GM, Huang H, Yan Q, Alvero R *et al.* Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. *N Engl J Med* 2014;371: 119-129.

Liu C, Feng G, Huang W, Wang Q, Yang S, Tan J, Fu J, Liu D. Comparison of clomiphene citrate and letrozole for ovulation induction in women with polycystic ovary syndrome: a prospective randomized trial. *Gynecological endocrinology* 2017: 1-5.

Lizneva D, Suturina L, Walker W, Brakta S, Gavrilova-Jordan L, Azziz R. Criteria, prevalence, and phenotypes of polycystic ovary syndrome. *Fertil Steril* 2016;106: 6-15.

Lopez E, Gunby J, Daya S, Parrilla JJ, Abad L, Balasch J. Ovulation induction in women with polycystic ovary syndrome: randomized trial of clomiphene citrate versus low-dose recombinant FSH as first line therapy. *Reproductive Biomedicine Online* 2004;9: 382-390.

Lord J, Thomas R, Fox B, Acharya U, Wilkin T. The effect of metformin on fat distribution and the metabolic syndrome in women with polycystic ovary syndrome--a randomised, double-blind, placebo-controlled trial. *BJOG: An International Journal of Obstetrics & Gynaecology* 2006;113: 817-824.

Lorzadeh N, Kazemirad S, Mohammadi Z. Comparison of effects letrozole and clomiphene citrate for ovulation induction in women with polycystic ovary syndrome [in Persian]. *Iranian Journal of Obstetrics, Gynecology and Infertility* 2011;14.

Maged AM, Elsawah H, Abdelhafez A, Bakry A, Mostafa WAI. The adjuvant effect of metformin and N-acetylcysteine to clomiphene citrate in induction of ovulation in patients with Polycystic Ovary Syndrome. *Gynecological endocrinology* 2015.

Mitwally MF, Casper RF. Use of an aromatase inhibitor for induction of ovulation in patients with an inadequate response to clomiphene citrate. *Fertil Steril* 2001;75: 305-309.

Mobusher I. Comparison of the efficacy of letrozole and clomiphene citrate for ovulation induction in infertile women with polycystic ovary syndrome. *Pakistan Journal of Medical and Health Sciences* 2014;8: 905-908.

Moll E, Bossuyt PM, Korevaar JC, Lambalk CB, van der Veen F. Effect of clomifene citrate plus metformin and clomifene citrate plus placebo on induction of ovulation in women with newly diagnosed polycystic ovary syndrome: randomised double blind clinical trial. *BMJ (Clinical research ed)* 2006;332: 1485.

Morin-Papunen L, Rantala AS, Unkila-Kallio L, Tiitinen A, Hippelainen M, Perheentupa A, Tinkanen H, Bloigu R, Puukka K, Ruokonen A *et al.* Metformin improves pregnancy and live-birth rates in women with polycystic ovary syndrome (PCOS): a multicenter, double-blind, placebo-controlled randomized trial. *Journal of Clinical Endocrinology & Metabolism* 2012;97: 1492-1500.

Morley LC, Tang T, Yasmin E, Norman RJ, Balen AH. Insulin-sensitising drugs (metformin, rosiglitazone, pioglitazone, D-chiro-inositol) for women with polycystic ovary syndrome, oligo amenorrhoea and subfertility. *Cochrane Database Syst Rev* 2017;11: CD003053.

Moussa AA, Torky H, Dief O, Elwahed AA, Senna HA. The effect of clomiphene citrate versus tamoxifen versus letrozol on endometrial thickness and blood flow in ovulation induction in women with polycystic ovaries. *Acta Medica International* 2016;3: 88-92.

Naderpoor N, Shorakae S, de Courten B, Misso ML, Moran LJ, Teede HJ. Metformin and lifestyle modification in polycystic ovary syndrome: systematic review and meta-analysis. *Hum Reprod Update* 2015;21: 560-574.

Nazik H, Kumtepe Y. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction for women with polycystic ovarian syndrome. *HealthMED* 2012;6: 879-883.

Palomba S, Orio F, Jr., Falbo A, Manguso F, Russo T, Cascella T, Tolino A, Carmina E, Colao A, Zullo F. Prospective parallel randomized, double-blind, double-dummy controlled clinical trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in nonobese anovulatory women with polycystic ovary syndrome. *Journal of Clinical Endocrinology & Metabolism* 2005;90: 4068-4074.

Riley RD, Jackson D, Salanti G, Burke DL, Price M, Kirkham J, White IR. Multivariate and network meta-analysis of multiple outcomes and multiple treatments: rationale, concepts, and examples. *BMJ* 2017;358: j3932.

Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340: c221.

Robinson R, Swezey M, Propst A, Bates G. Metformin added to clomiphene citrate does not improve pregnancy rates in hyperandrogenic, chronic anovulatory women: A randomized trial. *Fertility and sterility* 2003;80: S273-S274.

Rotterdam ESHRE/ASRM-Sponsored PCOS Consensus Workshop Group. Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS). *Hum Reprod* 2004;19: 41-47.

Roy K, Baruah J, Singla S, Sharma J, Singh N, Jain S, Goyal M. A prospective randomized trial comparing the efficacy of Letrozole and Clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. *Journal of Human Reproductive Sciences* 2012;5: 20-25.

Sahin Y, Yirmibes U, Kelestimur F, Aygen E. The effects of metformin on insulin resistance, clomiphene-induced ovulation and pregnancy rates in women with polycystic ovary syndrome. *European Journal of Obstetrics, Gynecology, & Reproductive Biology* 2004;113: 214-220.

Santonocito V, Rapisarda V, Abruzzo SRM, Pollicino R, Coco L, Zarbo G. Comparison between clomiphene citrate and metformin for induction of ovulatory cycles in infertile nonobese women with polycystic ovary syndrome [in Italian]. *Giornale Italiano di Ostetricia e Ginecologia* 2009;31: 455-460.

Selim MF, Borg TF. Letrozole and clomiphene citrate effect on endometrial and subendometrial vascularity in treating infertility in women with polycystic ovary syndrome. *Journal of Gynecologic Surgery* 2012;28: 405-410.

Seyedoshohadaei F, Zandvakily F, Shahgeibi S. Comparison of the effectiveness of clomiphene citrate, tamoxifen and letrozole in ovulation induction in infertility due to isolated unovulation. *Iranian Journal of Reproductive Medicine* 2012;10: 531-536.

Sharief M, Nafee NR. Comparison of letrazole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. *Journal of the Pakistan Medical Association* 2015;65: 1149-1152.

Sheikh-El-Arab Elsedeek M, Elmaghraby HAH. Predictors and characteristics of letrozole induced ovulation in comparison with clomiphene induced ovulation in anovulatory PCOS women. *Middle East Fertility Society Journal* 2011;16: 125-130.

Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF, Group P-ID. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA* 2015;313: 1657-1665.

Tang T, Glanville J, Hayden CJ, White D, Barth JH, Balen AH. Combined lifestyle modification and metformin in obese patients with polycystic ovary syndrome. A randomized, placebo-controlled, double-blind multicentre study. *Human Reproduction* 2006;21: 80-89.

Teede HJ, Misso ML, Costello MF, Dokras A, Laven J, Moran L, Piltonen T, Norman RJ, International PN. Recommendations from the international evidence-based guideline for the assessment and management of polycystic ovary syndrome. *Hum Reprod* 2018;33: 1602-1618.

Tierney JF, Pignon JP, Gueffyier F, Clarke M, Askie L, Vale CL, Burdett S, Cochrane IPDM-aMG. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. *J Clin Epidemiol* 2015;68: 1325-1335.

Topçu HO, Batioğlu AS, İslimye M. Tamoxifen versus clomiphene citrate for ovulation induction in women with polycystic ovary syndrome: A prospective randomized trial. *Journal of reproductive medicine* 2017;62: 507-512.

Vegetti W, Riccaboni A, Columbo M, Baroni E, Diaferia D, Ragni G, Crosignani PG. Randomized study of induction of ovulation by two different molecules with antioestrogenic effects, in patients with chronic anovulation disorders. 1999, pp. S234-235.

Wang R, Kim BV, van Wely M, Johnson NP, Costello MF, Zhang H, Ng EH, Legro RS, Bhattacharya S, Norman RJ *et al.* Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis. *BMJ* 2017;356: j138.

Wang R, Mol BW. The Rotterdam criteria for polycystic ovary syndrome: evidence-based criteria? *Hum Reprod* 2017;32: 261-264.

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

Williams CD, Pastore LM, Shelly WB, Bailey AP, Baras DC, Bateman BG. A randomized, placebo-controlled study of the influence of instant-release metformin on response to clomiphene citrate and time to conception in polycystic ovary syndrome. *Fertility and sterility* 2009;92: S105.

Wu XK, Stener-Victorin E, Kuang HY, Ma HL, Gao JS, Xie LZ, Hou LH, Hu ZX, Shao XG, Ge J *et al.* Effect of Acupuncture and Clomiphene in Chinese Women With Polycystic Ovary Syndrome: A Randomized Clinical Trial. *JAMA* 2017;317: 2502-2514.

Zain MM, Jamaluddin R, Ibrahim A, Norman RJ. Comparison of clomiphene citrate, metformin, or the combination of both for first-line ovulation induction, achievement of pregnancy, and live birth in Asian women with polycystic ovary syndrome: a randomized controlled trial. *Fertility & Sterility* 2009;91: 514-521.

Zeinalzadeh M, Basirat Z, Esmailpour M. Efficacy of letrozole in ovulation induction compared to that of clomiphene citrate in patients with polycystic ovarian syndrome. *Journal of reproductive medicine* 2010;55: 36-40.

Chapter 9

Conclusion

SUMMARY OF THE THESIS

In this thesis, we introduced network meta-analysis in reproductive medicine and apply this method to answer clinical questions on decision making among multiple treatment choices. We further used individual participant data (IPD) meta-analysis to provide insight into personalised treatment choice in reproductive medicine.

In **chapter 1**, we provided a general introduction on reproductive medicine, described the limitations of conventional evidence synthesis method and set forth objectives to overcome these barriers by using network meta-analysis and IPD meta-analysis in reproductive medicine.

In **chapter 2**, we highlighted the key assumptions in network meta-analyses and summarised existing network meta-analyses in reproductive medicine. We identified the challenges in the conduct and report of network meta-analyses.

In **chapter 3-6**, we used network meta-analyses according to the current standard methodological practice in a series of clinical scenarios where decisions need to be made among multiple treatments, including the first-line ovulation induction for WHO II anovulation (PCOS), second-line ovulation induction for PCOS, treatment strategies for unexplained infertility and the use of different contrast media during tubal patency tests to improve fertility outcomes. We identified top-ranking interventions for these clinical conditions and provided evidence for clinical practice and future research.

We then focused on the treatment choices of PCOS, addressing the need for IPD meta-analysis to guide personalised treatment choices and to refine the existing diagnostic criteria in **chapter 7**.

In **chapter 8**, we further illustrated an example of IPD meta-analysis on the first-line ovulation induction for PCOS, highlighting the importance of hyperandrogenemia and insulin resistance in guiding personalised clinical practice.

The evidence generated in this thesis is ready to implement into the development of evidence-based clinical guidelines and to guide daily clinical practice. The application of

these novel evidence synthesis methods extends the dimensions of evidence-based reproductive medicine by incorporating the comparisons of multiple treatments and personalised treatment strategies, which facilitates more patient-centred reproductive health care.

DIRECTIONS FOR FUTURE RESEARCH

Apart from the above-mentioned direct clinical implications, IPD meta-analyses and network meta-analyses are ideal for the development of priority lists that guide future research. IPD meta-analyses are useful to inform the design, conduct and analysis of new trials (Tierney et al., 2015). Based on treatment-covariate interactions observed in IPD meta-analyses, new trials should be encouraged to incorporate these treatment selection markers in their design to guide treatment decision (Janes et al., 2011). In network meta-analysis, only those top-ranking interventions should be prioritized for future research. For example, in our network meta-analyses for first-line ovulation induction in women with WHO II anovulation, letrozole alone and the combination of clomiphene citrate and metformin are considered the most effective treatments (Wang et al., 2017). These options will be evaluated in a two-by-two factorial randomised controlled trial recently funded by the National Institute for Health Research in the UK (Coomarasamy et al., 2019). This new trial will compare clomiphene and letrozole, with and without metformin in 1050 women with PCOS (Coomarasamy et al., 2019).

Evidence-based medicine is rapidly progressing over the past three decades, and the three principle tenets of evidence-based medicine are an increasingly sophisticated hierarchy of evidence, the need for systematic summaries of the best evidence to guide care, and the requirement for considering patient values in important clinical decisions (Djulbegovic and Guyatt 2017). Emerging next-generation evidence synthesis methods have been developed and implemented in the practice of evidence-based medicine, including network meta-analyses, IPD meta-analyses, prospective meta-analyses, umbrella reviews and living systematic reviews (Elliott et al., 2014; Ioannidis, 2017). In some cases, it may be beneficial to combine these methods. Individual-level data from each eligible trial can enable controlling for participant-level covariates, and thus improve consistency and precision of network meta-analysis, in particular when treatments are potentially influenced by participant-level covariates (Debray et al., 2018). In some cases, individual participant data

can enable network meta-analysis that would otherwise not be advisable due to violations of the transitivity assumption. In a prospective meta-analysis, trials are identified and determined to be eligible for inclusion before any results of the studies related to the prospective meta-analysis research question are known. For high-priority research questions with multiple available interventions, but insufficient previous evidence to conduct a network meta-analysis, the prospective planning of a number of studies comparing promising interventions as part of a prospective network meta-analysis could be considered (Seidler et al., 2019).

New evidence synthesis methods including network meta-analyses and IPD meta-analyses have limitations. Network meta-analyses do not involve new randomisation for the indirect comparisons and therefore they rely on the validity of transitivity assumption (Salanti, 2012). They could result in misleading results if the assumption is not considered carefully. IPD meta-analyses are time consuming, expensive and the validity of the results also rely on the availability of IPD (Broeze et al., 2010). The application of both methods to answer the clinical questions requires multidiscipline collaborations, and should be protocol based and clearly reported according to the current standard (Stewart at al., 2015; Hutton at al., 2015)

These novel evidence synthesis methods are promising in evidence-based reproductive medicine. These methods offer the opportunities with recommendations to improve the design, conduct and quality of future primary studies, which in return could improve the certainty of the overall evidence in systematic reviews. The evidence will ultimately improve the quality of care and couples' outcome in reproductive medicine.

REFERENCES

Broeze KA, Opmeer BC, van der Veen F, Bossuyt PM, Bhattacharya S, Mol BW. Individual patient data meta-analysis: a promising approach for evidence synthesis in reproductive medicine. *Hum Reprod Update*. 2010;16(6):561-7.

Coomarasamy A, Jayaprakasan K, Raine-Fenning N, Maheshwari A, Gallos I, Webber L, Choudhary M, Mathur R, Sunkara S, Khalaf Y, Devall A, Drakeley A, Middleton L, Feeney A, Pundir J, Balen A, Anderson R, Nelson S, Thangaratinam S, Bhattacharya S, Roberts T, Arlt W, Cheong Y. Letrozole or Clomifene, with or without metformin, for ovulation induction in women with polycystic ovary syndrome: a 2x2 factorial design randomised trial (The LOCI trial). *National Institute for Health Research Funding and Awards*. 2019. https://fundingawards.nihr.ac.uk/award/17/116/01

Debray TP, Schuit E, Efthimiou O, Reitsma JB, Ioannidis JP, Salanti G, Moons KG; GetReal Workpackage. An overview of methods for network meta-analysis using individual participant data: when do benefits arise? *Stat Methods Med Res*. 2018;27(5):1351-1364.

Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. *Lancet.* 2017;390: 415-423.

Elliott JH, Turner T, Clavisi O, Thomas J, Higgins JP, Mavergames C, Gruen RL. Living systematic reviews: an emerging opportunity to narrow the evidence-practice gap. *PLoS Med*. 2014 Feb 18;11(2):e1001603.

Hutton B, Salanti G, Caldwell DM, Chaimani A, Schmid CH, Cameron C, Ioannidis JP, Straus S, Thorlund K, Jansen JP, Mulrow C, Catalá-López F, Gøtzsche PC, Dickersin K, Boutron I, Altman DG, Moher D. *Ann Intern Med*. 2015;162(11):777-84.

The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations.

Janes H, Pepe MS, Bossuyt PM, Barlow WE. Measuring the performance of markers for guiding treatment decisions. *Ann Intern Med* 2011;154: 253-259.

Ioannidis J. Next-generation systematic reviews: prospective meta-analysis, individual-level data, networks and umbrella reviews. *Br J Sports Med*. 2017;51(20):1456-1458.

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Res Synth Methods*. 2012;3(2):80-97.

Seidler AL, Hunter K, Cheyne S, Ghersi D, Berlin J, Askie L. A guide to prospective meta-analysis. *BMJ*. 2019. In press.

Stewart LA, Clarke M, Rovers M, Riley RD, Simmonds M, Stewart G, Tierney JF; PRISMA-IPD Development Group. Preferred Reporting Items for Systematic Review and Meta-Analyses of individual participant data: the PRISMA-IPD Statement. *JAMA*. 2015;313(16):1657-65.

Tierney JF, Pignon JP, Gueffyier F, Clarke M, Askie L, Vale CL, Burdett S, Cochrane IPDM-aMG. How individual participant data meta-analyses have influenced trial design, conduct, and analysis. *J Clin Epidemiol* 2015;68: 1325-1335.

Wang R, Kim BV, van Wely M, Johnson NP, Costello MF, Zhang H, Ng EH, Legro RS, Bhattacharya S, Norman RJ *et al.* Treatment strategies for women with WHO group II anovulation: systematic review and network meta-analysis. *BMJ* 2017;356: j138.

Appendices

List of conference presentations

2019 The 9th Congress of the Asia Pacific Initiative on Reproduction (ASPIRE 2019), Hong Kong, 2-5 May 2019. Invited presentation.
What is an individual participant data meta-analysis?

2018 Fertility Society of Australia 2018 Annual Scientific Meeting (FSA 2018), Melbourne, Australia, 9-12 Sep 2018. Oral presentation.
Interventions for unexplained infertility: a Cochrane systematic review and network meta-analysis

2018 The 34th Annual Meeting of The European Society of Human Reproduction and Embryology (ESHRE 2018), Barcelona, Spain, 1-4 Jul, 2018. Oral presentation.
Interventions for unexplained infertility: a systematic review and network meta-analysis

2017 Fertility Society of Australia 2017 Annual Scientific Meeting (FSA 2017), Adelaide, Australia, 15-18 Oct 2017. Oral presentation.
Tubal flushing for subfertility - systematic review and network meta-analysis

2017 The 33rd Annual Meeting of The European Society of Human Reproduction and Embryology (ESHRE 2017), Geneva, Switzerland, 2-5 Jul, 2017. Poster presentation.
Water-based and oil-based tubal flushing for subfertility, alone, both or none? Systematic review and network meta-analysis

2017 The 7th Congress of the Asia Pacific Initiative on Reproduction (ASPIRE 2017), Kuala Lumpur, Malaysia, 30 Mar - 2 Apr. Poster presentation.
Tailored first-line ovulation induction for women with PCOS - an individual participant data (IPD) meta-analysis

2017 Fertility Conference 2017 (The Association of Clinical Embryologists, British Fertility Society and the Society for Reproduction & Fertility annual meeting), Edinburgh, UK, 5-7 Jan 2017. Invited presentation.
Personalized first-line ovulation induction for women with PCOS – an individual participant data (IPD) meta-analysis

2017 The 14th Annual Meeting of AE-PCOS Society, Lorne, Victoria, Australia, 10-12 Nov, 2016. Poster presentation.

Treatment strategies for women with WHO II anovulation - a systematic review and network meta-analysis

2016 Fertility Society of Australia 2016 Annual Scientific Meeting (FSA 2016), Perth, Australia. 4-7 Sep, 2016. Oral presentation.

Clomiphene, metformin, letrozole, tamoxifen or combined clomiphene-metformin for polycystic ovary syndrome – a systematic review and individual participant data network meta-analysis

List of workshop attendance

2017 Writing a systematic review following Cochrane methods workshop. Cochrane Australia.

8-10 May 2017, Adelaide, Australia.

2018 Systematic reviews and meta-analysis of IPD. Utrecht University.

25-29 June 2018, Utrecht, The Netherlands.

2018 How to understand, appraise and write a network meta-analysis. Neuroscience Research Australia.

2-4 Nov 2018, Sydney, Australia

2019 Systematic Reviews & Meta-analysis of Prognosis Studies. NHMRC Clinical Trials Centre.

19-21 June 2019, Sydney, Australia

List of other publications during PhD candidature

1. Li J, Wu Q, Wang CC, **Wang R**, Ng EHY, Liu JP, Mol BWJ, Wu XK, Li WT; PCOSAct Study Group. Endocrine characteristics, body mass index and metabolic syndrome in women with polycystic ovary syndrome. *Reprod Biomed Online*. 2019. pii: S1472-6483(19)30601-7
2. Dang VQ, Vuong LN, Ho TM, Ha AN, Nguyen QN, Truong BT, Pham QT, **Wang R**, Norman RJ, Mol BW. The effectiveness of ICSI versus conventional IVF in couples with non-male factor infertility: study protocol for a randomised controlled trial. *Hum Reprod Open*. 2019;2019(2):hoz006.
3. Wohling J, Edge N, Pena-Leal D, **Wang R**, Mol BW, Dekker G. Clinical and financial evaluation of carbetocin as postpartum haemorrhage prophylaxis at caesarean section: A retrospective cohort study. *Aust N Z J Obstet Gynaecol*. 2018. doi: 10.1111/ajo.12907.
4. Kemper JM, **Wang R**, Vuong LN, Mol BW. Single embryo transfer with frozen transfer of all remaining embryos without further embryonic testing should be the standard of care in IVF. *BJOG*. 2019;126(2):142-144.
5. Vuong LN, Ho VNA, Ho TM, Dang VQ, Phung TH, Giang NH, Le AH, Pham TD, **Wang R**, Norman RJ, Smitz J, Gilchrist RB, Mol BW. Effectiveness and safety of in vitro maturation of oocytes versus in vitro fertilisation in women with high antral follicle count: study protocol for a randomised controlled trial. *BMJ Open*. 2018 Dec 9;8(12):e023413.
6. Huang S, Du X, **Wang R**, Li R, Wang H, Luo L, O'Leary S, Qiao J, Mol BWJ. Ovulation induction and intrauterine insemination in infertile women with polycystic ovary syndrome: A comparison of drugs. *Eur J Obstet Gynecol Reprod Biol*. 2018;231:117-121.
7. van Rijswijk J, Pham CT, Dreyer K, Verhoeve HR, Hoek A, de Bruin JP, Nap AW, **Wang R**, Lambalk CB, Hompes PGA, Mijatovic V, Karnon JD, Mol BW. Oil-based or water-based contrast for hysterosalpingography in infertile women: a cost-effective analysis of a randomized controlled trial. *Fertil Steril*. 2018;110(4):754-760
8. Duffy JMN, Bhattacharya S, Curtis C, Evers JLH, Farquharson RG, Franik S, Khalaf Y, Legro RS, Lensen S, Mol BW, Niederberger C, Ng EHY, Repping S, Strandell A, Torrance HL, Vail A, van Wely M, Vuong NL, Wang AY, **Wang R**, Wilkinson J, Youssef MA, Farquhar CM, COMMIT: Core Outcomes Measures for Infertility Trials. A protocol developing, disseminating and implementing a core outcome set for infertility. *Hum Reprod Open.* 2018;2018(3):hoy007.

9. 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. *Fertil Steril*. 2018;109(5):872-878
10. Kwan I, **Wang R**, Pearce E, Bhattacharya S. Pain relief for women undergoing oocyte retrieval for assisted reproduction. *Cochrane Database Syst Rev*. 2018;5:CD004829. doi: 10.1002/14651858.CD004829.pub4.
11. **Wang R**, Chien P, Leung TY. Is it time to move towards non-invasive prenatal screening for Down syndrome and screening for pre-eclampsia? *BJOG*. 2018;125(4):397-398.
12. Gadalla MA, **Wang R**, van Wely M, Mol BWJ. How should we report outcomes in reproductive medicine? *Ultrasound Obstet Gynecol*. 2018;51(1):7-9.
13. 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.
14. Coolen AWM, Bui BN, Dietz V, **Wang R**, van Montfoort APA, Mol BWJ, et al. The treatment of post-hysterectomy vaginal vault prolapse: a systematic review and meta-analysis. *Int Urogynecol J*. 2017. *Int Urogynecol J*. 2017;28(12):1767-1783
15. Tjon-Kon-Fat RI, **Wang R**, Eijkemans MJC, Bossuyt PMM, Mol BWJ, van der Veen F, et al. Interventions for unexplained subfertility: a systematic review and network meta-analysis (Protocol). Cochrane Database Syst Rev. 2017;6. DOI: 10.1002/14651858.CD012692.
16. Li W, Lin X, **Wang R**, Wang F, Xie S, Tse LA. Hormone therapy and lung cancer mortality in women: Systematic review and meta-analysis. *Steroids*. 2017;118:47-54.

List of co-authors and affiliations

Saad A Amer	Department of Obstetrics and Gynaecology, University of Nottingham, Royal Derby Hospital, Derby, UK
Lisa Askie	NHMRC Clinical Trials Centre, University of Sydney, Sydney, Australia
Siladitya Bhattacharya	Aberdeen Centre for Women's Health Research, University of Aberdeen, Aberdeen, UK
Esmée M Bordewijk	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Patrick M M Bossuyt	Department of Clinical Epidemiology, Biostatistics and Bioinformatics, UMC Amsterdam, University of Amsterdam, Amsterdam, the Netherlands
Michael F Costello	School of Women's and Children's Health, University of New South Wales, Sydney, Australia
Nora A Danhof	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Kim Dreyer	Department of Reproductive Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
Marinus J C Eijkemans	Department of Biostatistics and Research Support, Julius Center, UMC Utrecht, Utrecht, the Netherlands
Angela Falbo	Department of Obstetrics and Gynecology, Azienda USL-IRCCS di Reggio Emilia, Italy
Grasso Federica	Department of Surgery Obstetrics and Gynecology, University of Catania, Catania, Italy
Jingshu Gao	Department of Obstetrics and Gynecology, First Affiliated Hospital, Heilongjiang University of Chinese Medicine, Harbin, China
Roy Homburg	Homerton Fertility Centre, Homerton University Hospital, London, UK

Wei Huang	Department of Obstetrics and Gynecology, West China Second University Hospital of Sichuan University, Chengdu, China
Neil P Johnson	Robinson Research Institute and Adelaide Medical School, University of Adelaide, North Adelaide, SA, Australia; Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand
Esther R Jones	Robinson Research Institute and Adelaide Medical School, University of Adelaide, North Adelaide, SA, Australia
Sujata Kar	Department of Obstetrics and Gynaecology, Kar Clinic and Hospital, Bhubaneswar, India
Bobae V Kim	Robinson Research Institute and Adelaide Medical School, University of Adelaide, North Adelaide, SA, Australia
Tamar E König	Department of Obstetrics and Gynecology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
Richard S Legro	Department of Obstetrics and Gynecology, Pennsylvania State University, Hershey, Pennsylvania, USA
Wentao Li	Department of Obstetrics and Gynaecology, Monash University, Clayton, VIC, Australia
Jonathan Lord	Department of Obstetrics and Gynaecology, Royal Cornwall Hospital, University of Exeter Medical School, Truro, UK
Velja Mijatovic	Department of Reproductive Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
Monique H Mochtar	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Ben Willem Mol	Robinson Research Institute and Adelaide Medical School, University of Adelaide, North Adelaide, SA, Australia; Department of Obstetrics and Gynaecology, Monash University, Clayton, VIC, Australia

Etelka Moll	Department of Obstetrics and Gynaecology, Onze Lieve Vrouwe Gasthuis (OLVG) West, Amsterdam, the Netherlands
Laure Morin-Papunen	Department of Obstetrics and Gynecology, University of Oulu and Oulu University Hospital, Medical Research Center, PEDEGO Research Unit, Oulu, Finland
Hakan Nazik	Department of Obstetrics and Gynaecology, Adana City Training and Research Hospital, Adana, Turkey
Ernest Hung Yu Ng	Department of Obstetrics and Gynaecology, University of Hong Kong, Hong Kong SAR, China
Robert J Norman	Robinson Research Institute and Adelaide Medical School, University of Adelaide, North Adelaide, SA, Australia; FertilitySA, Adelaide, SA, Australia
Ülkü Özmen	Department of Obstetrics and Gynecology, School of Medicine, Zonguldak Bulent Ecevit University, Kozlu, Zonguldak, Turkey
Stefano Palomba	Department of Obstetrics and Gynecology, Grande Ospedale Metropolitano of Reggio Calabria, Reggio Calabria, Italy
Yilmaz Sahin	Department of Obstetrics and Gynecology, Faculty of Medicine, Erciyes University, Kayseri, Turkey
Anna Lene Seidler	NHMRC Clinical Trials Centre, University of Sydney, Sydney, Australia
Raissa I Tjon-Kon-Fat	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Fulco van der Veen	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Joukje van Rijswijk	Department of Reproductive Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands
Nienke van Welie	Department of Reproductive Medicine, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands

Madelon van Wely	Centre for Reproductive Medicine, Amsterdam UMC, University of Amsterdam, Amsterdam, the Netherlands
Walter Vegetti	Infertility Unit, Fondazione I.R.C.C.S. Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy
Christopher D Williams	Reproductive Medicine and Surgery Center of Virginia, Charlottesville, Virginia, USA
Xiaoke Wu	Department of Obstetrics and Gynecology, First Affiliated Hospital, Heilongjiang University of Chinese Medicine, Harbin, China
Heping Zhang	Department of Biostatistics, Yale University School of Public Health, New Haven, Connecticut, USA
Hanwang Zhang	Reproductive Medicine Centre, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Supplemental materials for Chapter 2

Supplemental Table 1 Search strategies 6/8/2019

Data source	Search strategy
PubMed	("network meta-analysis"[MeSH] OR network meta-analys* OR ((mixed treatment* OR multiple treatment* OR mixed comparison* OR indirect comparison*) AND meta-analys*)) AND (fertil* OR infertil* OR pregnan* OR subfertil* OR live birth*)

Supplemental materials for Chapter 3
Appendix 1 Search strategies
1a. MEDLINE search strategy Database: Ovid MEDLINE
1 exp Polycystic Ovary Syndrome/
2 Polycystic Ovar$.tw.
3 PCOS.tw.
4 PCOD.tw.
5 PCO.tw.
6 (stein-leventhal or leventhal).tw.
7 (ovar$ adj (scelerocystic or polycystic or degeneration)).tw.
8 anovulat$.ti,ab,sh,tw.
9 oligo ovulat$.ti,ab,sh,tw.
10 or/1-9
11 randomized controlled trial.pt.
12 controlled clinical trial.pt.
13 randomly.ab,ti.
14 randomized.ab,ti.
15 (crossover or cross over).tw.
16 placebo.tw.
17 RCT.tw.
18 trial.ti.
19 clinical trials as topic.sh.
20 or/11-19
21 exp animals/ not humans.sh.
22 20 not 21
23 fertil$.ti,ab,sh,tw.
24 infertil$.ti,ab,sh,tw.
25 subfertil$.ti,ab,sh,tw.
26 pregnan$.ti,ab,sh,tw.
27 exp ovulation induction/ or exp superovulation/
28 (ovulat$ adj2 induc$).tw.
29 (ovar$ adj2 stimulat$).tw.
30 superovulat$.tw.
31 or/23-30
32 10 and 22 and 31

1b. Embase search strategy Database: EMBASE.com
#1 'ovary polycystic disease'/exp OR 'stein leventhal syndrome'/exp
#2 (polycystic NEAR/2 ovar*):de,ab,ti
#3 pcos:de,ab,ti OR pcod:de,ab,ti OR pco:de,ab,ti
#4 leventhal:de,ab,ti
#5 (ovar* NEAR/2 (scelerocystic OR degeneration)):de,ab,ti
#6 'anovulation'/exp
#7 anovulat*:de,ab,ti
#8 (oligo NEAR/2 ovulat*):de,ab,ti
#9 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10 'randomized controlled trial'/exp
#11 'controlled clinical trial'/exp
#12 randomized:de,ab,ti
#13 randomly:de,ab,ti
#14 trial:ti
#15 plecebo:de,ab,ti
#16 rct:de,ab,ti
#17 crossover:de,ab,ti OR (cross NEAR/1 over):de,ab,ti
#18 'clinical trial' OR 'clinical trials':de
#19 #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18
#20 #19 AND [animals]/lim NOT [humans]/lim
#21 #19 NOT #20
#22'infertility'/exp OR 'fertility'/exp OR 'subfertility'/exp

#23 infertil*:de,ab,ti OR subfertil*:de,ab,ti OR feril*:de,ab,ti
#24 pregnan*:de,ab,ti
#25 'pregnancy'/exp
#26 'ovulation induction'/exp OR 'superovulation'/exp
#27 (ovulat* NEAR/2 induc*):de,ab,ti
#28 (ovar* NEAR/2 stimulat*):de,ab,ti
#29 superovulat*:de,ab,ti
#30 #22 OR #23 OR #24 OR #25 OR #26 OR #27 OR #28 OR #29
#31 #9 AND #21 AND #30

1c. Database: EBM Reviews - Cochrane Central Register of Controlled Trials

#1	[mh "Polycystic Ovary Syndrome"]
#2	(polycystic near ovar*):kw,ab,ti
#3	pcos:kw,ab,ti or pcod:kw,ab,ti or pco:kw,ab,ti
#4	leventhal:kw,ab,ti
#5	(ovar* near (scelerocystic or degeneration)):kw,ab,ti
#6	anovulat*:kw,ab,ti
#7	oligo near ovulat*:kw,ab,ti
#8	[mh anovulation]
#9	#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8
#10	randomized controlled trial:pt
#11	controlled clinical trial:pt
#12	plecebo:kw,ti,ab
#13	randomly:kw,ti,ab
#14	RCT:kw,ti,ab
#15	trial:ti
#16	crossover:kw,ti,ab or (cross next over):kw,ti,ab
#17	#10 or #11 or #12 or #13 or #14 or #15 or #16
#18	[mh infertility]
#19	[mh fertility]
#20	[mh pregnancy]
#21	infertil*:kw,ti,ab
#22	fertil*:kw,ti,ab
#23	subfertil*:kw,ti,ab
#24	pregnan*:kw,ti,ab
#25	[mh "Ovulation Induction"] or [mh superovulation]
#26	ovulat* near induc*:kw,ti,ab
#27	ovar* near stimulat*:kw,ti,ab
#28	superovulat*:kw,ti,ab
#29	#18 or #19 or #20 or #21 or #22 or #23 or #24 or #25 or #26 or #27 or #28
#30	#9 and #17 and #29

Appendix 2 Characteristics of included studies

Study	Interventions	Age (mean)	BMI (mean)	DOI (mean years)	Inclusion criteria	Sample Size	Previous Treatment	Country	Setting	Maximum of treatment cycles	IUI or TI
Abuelghar 2013[1]	CC MF+CC	28.4 27.6	28.1 28.6	2.8 3.2	Overweight and obese infertile women with PCOS (Rotterdam criteria)	66	unknown	Egypt	single-centre	1	TI
Amer 2009[2]	CC LOD	29.1 28.1	26.1 26.2	1.8 2.1	PCOS (at least 2 of the following 3 features: clinical [oligo/amenorrhoea and/or Hyperandrogenaemia], biochemical [LH≥10 IU/l, LH/FSH ratio ≥2, testosterone>2.6 nmol/l or free androgen index (FAI) >5] and/or sonographic (polycystic ovaries) features.)	72	naive	UK	single-centre	6	TI
Amer 2015[3]	CC LET	NA	NA	NA	anovulatory women with PCOS	159	naive	UK	single-centre	7	TI
Atay 2006[4]	CC LET	26.2 27.1	25.8 26.1	2.4 2.2	Women with primary infertility and PCOS(oligo- or amenorrhoea and ovaries with at least 10 subcapsular cysts 2 – 10 mm in diameter and hyperechogenic stroma.)	106	unknown	Turkey	N/A	1	TI
Ayaz 2013[5]	CC MF+CC	31.3 32.0	NA[a]	NA	PCOS (the presence of two of the three following criteria:1. Polycystic ovaries [either 12 or more peripheral follicles or increased ovarian volume, > 10 cm^3]. 2. Oligo or anovulation [irregular cycles, amenorrhea]. 3. Clinical and/or biochemical signs of hyperandrogenism	42	unknown	Saudi Arabia	single-centre	6	TI

					[Acne, hirsutism, voice changes, and Clitoromegaly].)						
Aygen 2007[6]	CC LET	23.4 26.8	27.6 26.9	4.2 5.8	Infertility and PCOS (Rotterdam criteria)	10	unknown	Turkey	single-centre	6	TI
Badawy 2009[7]	CC LET	29.3 27.1	27.1 28.1	NA	Infertile women with PCOS (Rotterdam criteria)	438	unknown	Egypt	multi-centre	>1	TI
Badawy 2011[8]	CC TAM	25.8 26.2	29.9 30.5	1.5 1.4	PCOS (Rotterdam criteria)	371	unknown	Egypt	multi-centre	1	TI
Basirat 2012[9]	CC MF+CC	25.3 24.9	25.4 26.3	2.7 2.4	Infertile PCOS (Rotterdam criteria)	334	unknown	Iran	multi-centre	3	IUI
Bayar 2006[10]	CC LET	30.6 32.2	NA	3 5	anovulatory PCOS (Rotterdam criteria)	80	naive	Turkey	single-centre	>1	TI
Beigi 2006[11]	CC MF	NA	NA	NA	PCOS based on a history of hyperandrogenism, anovulation, oligomenorrhea or amenorrhea, diagnostic ultrasound and laboratory findings	70	unknown	Iran	single-centre	6	TI
Boonstanfar 2001[12]	CC TAM	26.5 26.6	30.2 30.9	3.7 3.5	anovulatory women with infertility	95	naive	USA	single-centre	>1	TI
Boudhraa 2010[13]	CC MF+CC	30.7 30.6	29.8 30.0	2.5[b]	PCOS (Rotterdam criteria) with subfertility	63	unknown	Tunis	single-centre	3-6	TI
Cudmore 1966[14]	CC PB	24.6 24.6	NA	NA	A diagnosis of secondary amenorrhea of at least 2 year's duration; persistent oligomenorrhea with no more than 4 periods in 1 year; or anovulatory infertility (infertility of more than 2 years' duration in which anovulation was the only cause found)	22	unknown	Canada	single-centre	3	TI
Dasari 2009[15]	CC MF+CC	NA[c]	NA[d]	NA	Infertile PCOS (Rotterdam criteria)	40	unknown	India	single-centre	6	TI
Dehbashi 2009[16]	CC LET	24.3 23.6	27.1 27.5	2.3 2.0	PCOS (Rotterdam criteria)	100	naive	Iran	single-centre	1	TI
El-Biely 2001[17]	CC MF+CC	25.7 26.4	27.4 28.7	4.7 4.5	Infertile obese patients with PCOS (oligomenorrhoea,	90	unknown	Egypt	single-centre	6	TI

					ultrasound findings of ≥ 10 ovarian cysts measuring 2-8mm around a dense stroma)						
Fleming 2002[18]	MF PB	28.6 29.2	34.2 35.0	NA	Women with oligomenorrhea or amenorrhea and PCO	42	naive	UK	single-centre	4	TI
Garcia 1985[19]	CC PB	27.6[e]	NA	NA	Anovulatory infertile women	49	unknown	USA	single-centre	5	TI
Homburg 2012[20]	CC FSH	29.4 29.8	25.7 25.1	2.1 2.1	anovulatory or oligo-ovulatory infertile women with PCOS (Rotterdam criteria)	302	naive	Netherlands, UK, Malta, Belgium, Argentina Colombia	multi-centre	3	TI/IUI
Jahan 2015[21]	CC LET MF	NA	NA	NA	PCOS	460	naive	Bangladesh	single-centre	6	TI
Johnson 1966[22]	CC PB	NA	NA	NA	Anovulatory women	65	mixed	USA	single-centre	1	TI
Johnson 2010A[23]	MF PB	29.5 29.2	38.0 37.6	3.3(2.4-5.9)[f] 3.4(2-5)[f]	anovulatory or oligo-ovulatory women with PCOS (Rotterdam criteria), BMI>32 kg/m^2	65	mixed	New Zealand	multi-centre	6	TI
Johnson 2010B[23]	CC MF MF+CC	28.2 28.9 29.2	26.2 26.5 26.9	2(1-3)[f] 1.5(1-4)[f] 2(1.5-5)[f]	anovulatory or oligo-ovulatory women with PCOS (Rotterdam criteria), BMI≤32 kg/m^2	106	mixed	New Zealand	multi-centre	6	TI
Kar 2012[24]	CC LET	26.3 26.3	26.0 25.9	3.1 3.1	infertile PCOS (Rotterdam criteria)	103	naive	India	single-centre	1	TI/IUI
Kar 2015[25]	CC MF MF+CC	25.8 25.2 26.6	26.5 24.5 27.2	2.8 1.7 2.5	PCOS (Rotterdam criteria), with the primary complaints of infertility and oligomenorrhea	105	naive	India	single-centre	6	TI
Karimzadeh 2007[26]	MF PB	27.2 28.6	28.8 29.5	5.6 6.2	PCOS (Rotterdam criteria)	200	unknown	Iran	single-centre	3	TI
Karimzadeh 2010[27]	CC MF MF+CC	27.5 27.3 27.3	27.2 27.2 28.0	4.1 3.9 4.6	infertile PCOS (Rotterdam criteria)	268	unknown	Iran	single-centre	6	TI
Keikha 2011[28]	CC	27.1	NA	2.9	infertile PCOS	116	naive	Iran	single-	1	TI

	LET	27.6		3.0					centre		
Khorram 2006[29]	CC MF+CC	28.0 28.4	38.8 35.3	NA	PCOS (anovulatory or oligo-ovulatory cycles, polycystic ovaries on a baseline ultrasound, hyperandrogenism) and infertility	31	naive	USA	single-centre	1	TI
Leanza 2014[30]	CC MF+CC	26-34[g]	NA	NA	PCOS (typical ultrasound situation, oligomenorrhea/amenorrhea, hyperandrogenism) with above 3 years of infertility, BMI>27.5	56	naive	Italy	single-centre	3	IUI
Legro 2007[31]	CC MF MF+CC	27.9 28.1 28.3	36.0 35.6 34.2	3.5 3.3 3.4	infertile women PCOS (oligomenorrhea and hyperandrogenemia)	626	mixed	USA	multi-centre	6	TI
Legro 2014[32]	CC LET	28.8 28.9	35.1 35.2	3.5 3.4	infertile women PCOS (Rotterdam criteria)	750	mixed	USA	multi-centre	5	TI
Liu 2015[33]	CC LET	NA	NA	NA	PCOS patients who have conception desire	134	unknown	China	single-centre	>1	TI
López 2004[34]	CC FSH	29(23-38)[f] 30(22-39)[f]	22.3 21.9	3(1-8)[f] 3(1-8)[f]	anovulatory infertility due to PCOS (Rotterdam criteria)	76	naive	Spain	single-centre	3	TI
Lord 2006[35]	MF PB	27.8 30.6	33.7 36.4	NA	PCOS (anovulation and a raised free androgen index (FAI) >5.0)	44	unknown	UK	single-centre	3	TI
Lorzadeh 2011[36]	CC LET	26.1 28.2	25.4 24.2	NA	PCOS (based on the chronic anovulation and clinical/lab-based hyperandrogenism), age <35, No successful pregnancy after one year of weekly (2-3 times) sexual contact without contraception.	100	unknown	Iran	single-centre	>1	TI
Maged 2015[37]	CC MF+CC	26.0 25.8	27.3 27.7	2.8 2.8	PCOS (Rotterdam criteria)	80	unknown	Egypt	single-centre	3	TI
Mobusher 2014[38]	CC LET	24.3 24.3	25.9 25.9	3.1 3.2	PCOS (Rotterdam criteria) and infertility	100	naive	Pakistan	single-centre	1	TI

Moll 2006[39]	CC MF+CC	28.4 27.9	27.8 28.5	1.3 1.6	PCOS (Rotterdam criteria), all women with chronic anovulation and polycystic ovaries diagnosed by transvaginal ultrasonography	225	naive	Netherlands	multi-centre	6	TI
Nazik 2012[40]	CC LET	27.8 25.6	25.9 24.7	4.4 3.4	PCOS (Rotterdam criteria)	64	naive	Turkey	single-centre	>1	TI
Palomba 2005[41]	CC MF	25.9 26.4	26.7 27.0	1.7 1.6	primary infertile anovulatory women with PCOS (NIH criteria)	100	naive	Italy	single-centre	6	TI
Raja 2005[42]	CC MF+CC	26.9 26.5	NA	4.9 4.2	Infertility and PCOS (the presence of polycystic ovaries on ultrasonography with two or more of the following criteria: Oligomenorrhoea [<6 cycles in preceding year); hirsutism; hyperandrogenism; Elevated LH or LH: FSH >2])	100	unknown	Pakistan	single-centre	6	TI
Ray 2012[43]	CC LET	29(20-35) [f] 28(19-35) [f]	28.5(24.2-33.6) [f] 28.8(23.2-34.6) [f]	2.4 2.2	Infertile PCOS (Rotterdam criteria)	147	unknown	India	single-centre	>1	TI
Robinson 2003[44]	CC MF+CC	NA	NA	NA	Women with a one-year history of infertility and diagnosed with hyperandrogenic oligoovulatory or anovulatory cycles as the sole etiology for their infertility	48	unknown	USA	single-centre	6	TI
Roy 2012[45]	CC LET	26.5 26.1	25.4 25.8	5.8 6.4	infertility and anovulatory PCOS (Rotterdam criteria), BMI<28	212	unknown	India	single-centre	3	TI
Sahin 2004[46]	CC MF+CC	24.5(19-28) [f] 27(21-31) [f]	25.7(23.1-35.7) [f] 30.4(24.6-33.9) [f]	3.5(1-8) [f] 5(2-10) [f]	Primary infertility and PCOS (on the basis of three or more of the following criteria: polycystic ovaries	21	unknown	Turkey	single-centre	6	TI

					on pelvic ultrasound examination, oligo/amenorrhoea, hirsutism, hyperandrogenaemia (total testosterone > 80 ng/dl and/or free testosterone > 3.18 pg/ml)) and elevated serum LH:FSH ratio (LH:FSH > 2))						
Santonocito 2009[47]	CC MF	27.4 28.1	27.1 26.8	1.7 1.6	infertility and anovulatory PCOS (Rotterdam criteria), BMI< 30 kg/m^2	36	unknown	Italy	single-centre	6	TI
Selim 2012[48]	CC LET	25.1 26.0	23.8 24.4	2.6 2.9	Infertile women with PCOS (Rotterdam criteria)	220	naive	Egypt	single-centre	1	TI
Seyedoshohadaei 2012[49]	CC LET TAM	24.7 26.9 25.4	NA	3.0 4.1 3.0	non-PCOS anovulatory infertility, and ovary without evidence of polycystic ovaries	150	unknown	Iran	single-centre	6	TI
Sharief 2015[50]	CC LET	25.3 26.1	27.8 28.1	2.3 2.4	primary infertility and anovulation due to PCOS (ultrasonographic polycystic ovaries plus one or more of the following: oligomenorrhoea, positive progesterone, withdrawal bleeding, hirsutism/acne, obesity, and Luteinizing hormone/Follicle-stimulating hormone (LH/FSH) ratio >2 or raised circulating androgen, normal thyroid stimulating hormone)	75	unknown	Iraq	single-centre	6	TI
Sh-El-Arab Elsedeek 2011[51]	CC LET	25.0 25.0	29.2 27.7	NA	Nulliparous PCOS (Rotterdam criteria), BMI ≤35	124	unknown	Egypt	single-centre	1	TI
Tang 2006[52]	MF PB	29.7 29.8	37.6 38.9	4.5 4.9	anovulatory PCOS (polycystic ovaries on	143	naive	UK	multi-centre	6	TI

					transvaginal scan, together with either oligomenorrhoea or amenorrhoea) and a BMI of >30,						
Vegetti 1999[53]	CC TAM	NA	NA	NA	Infertility and normogonadotropic anovulation	95	naive	Italy	single-centre	>1	TI
Williams 2009[54]	CC MF+CC	NA	NA	NA	women with PCOS who are attempting to conceive.	55	unknown	USA	N/A	6	TI
Zain 2009[55]	CC MF MF+CC	29.6 27.8 29.3	32.9 33.9 33.0	2.9 3.1 3.3	PCOS (Rotterdam criteria)	124	naive	Malaysia	single-centre	6	TI
Zeinalzadeh 2010[56]	CC LET	23.1 23.8	NA	2.6 2.4	PCOS (based on ultrasonography finding, oligomenorrhea and an increased LH/FSH ratio (>3))	107	naive	Iran	single-centre	1	IUI

(Abbreviations: CC, clomiphene citrate; PB, placebo or no treatment; LET, letrozole; MF, metformin; TAM, tamoxifen; FSH, follicle stimulating hormone; LOD, laparoscopic ovarian drilling; NA, not available; BMI, body mass index; DOI: Duration of infertility)

a. The percentages of women with BMI>25 in CC and CC+MF group are 71.4% and 56.7%, respectively.

b. The mean duration of infertility of all the participants (including both groups).

c. The percentages of women with age >31, 26-30 and 20-25 years are 8.3%, 41.7%, 50% in CC group and 18.8%, 43.8% and 37.5% in CC+MF group.

d. The percentages of women with BMI >25 and BMI < 25 are 37.5% and 62.5%, respectively.

e. in treatment group only

f. median (range)

g. range

Appendix 3 List of included studies

1. Abuelghar WM, Elkady OS, Khamees AA. Clomiphene citrate alone, in combination with metformin or in combination with pioglitazone as first line therapy in induction of ovulation in infertile women with polycystic ovary syndrome, a randomized controlled trial. Middle East Fertility Society Journal 2013;18(3):135-41
2. Amer S, Fakis A, Smith J, et al. Double blind cross-over randomized controlled trial comparing letrozole versus clomiphene citrate for ovulation induction in women with polycystic ovarian syndrome. Human Reproduction 2015;30:i96
3. Amer SA, Li TC, Metwally M, et al. Randomized controlled trial comparing laparoscopic ovarian diathermy with clomiphene citrate as a first-line method of ovulation induction in women with polycystic ovary syndrome. Human Reproduction 2009;24(1):219-25
4. Atay V, Cam C, Muhcu M, et al. Comparison of letrozole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. Journal of international medical research 2006;34(1):73-6
5. Ayaz A, Alwan Y, Farooq MU. Metformin-clomiphene citrate vs. clomiphene citrate alone: Polycystic ovarian syndrome. Journal of human reproductive sciences 2013;6(1):15-18
6. Aygen EM, Güzel Z, Özgün T, et al. The use of letrozole for ovulation induction in infertile women with polycystic ovarian syndrome. Erciyes Tip Dergisi 2007;29(3):195-200
7. Badawy A, Abdel Aal I, Abulatta M. Clomiphene citrate or letrozole for ovulation induction in women with polycystic ovarian syndrome: a prospective randomized trial. Fertility & Sterility 2009;92(3):849-52
8. Badawy A, Gibreal A. Clomiphene citrate versus tamoxifen for ovulation induction in women with PCOS: a prospective randomized trial. European Journal of Obstetrics, Gynecology, & Reproductive Biology 2011;159(1):151-4 doi: http://dx.doi.org/10.1016/j.ejogrb.2011.07.015.
9. Basirat Z, Kashifard M, Amiri MG. Enhanced ovarian folliclular development by Metformin does not correlate with pregnancy rate: A randomized trial. International Journal of Fertility and Sterility 2012;6(1):31-36
10. Bayar U, Basaran M, Kiran S, et al. Use of an aromatase inhibitor in patients with polycystic ovary syndrome: a prospective randomized trial. Fertility & Sterility 2006;86(5):1447-51
11. Beigi A. Randomized trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in polycystic ovary syndrome. Human reproduction (Oxford, England) 2006;21(Suppl):i129
12. Boostanfar R, Jain JK, Mishell DR, Jr., et al. A prospective randomized trial comparing clomiphene citrate with tamoxifen citrate for ovulation induction. Fertility & Sterility 2001;75(5):1024-6
13. Boudhraa K, Jellouli MA, Amri M, et al. [Indication of metformin in the management of hormonal dysfunction secondary to polycystic ovarian syndrome: prospective comparative study of 63 cases]. Tunisie medicale 2010;88(5):335-40
14. Cudmore DW, Tupper WR. Induction of ovulation with clomiphene citrate. A double-blind study. Fertil Steril 1966;17(3):363-73
15. Dasari P, Pranahita GK. The efficacy of metformin and clomiphene citrate combination compared with clomiphene citrate alone for ovulation induction in infertile patients with PCOS. Journal of human reproductive sciences 2009;2(1):18-22
16. Dehbashi S, Dehbashi S, Kazerooni T, et al. Comparison of the effects of letrozole and clomiphene citrate on ovulation and pregnancy rate in patients with polycystic ovary syndrome. Iranian Journal of Medical Sciences 2009;34(1):23-28
17. El-Biely MM, Habba M. The use of metformin to augment the induction of ovulation in obese infertile patients with polycystic ovary syndrome. Middle East Fertility Society Journal 2001;6(1):43-49
18. Fleming R, Hopkinson ZE, Wallace AM, et al. Ovarian function and metabolic factors in women with oligomenorrhea treated with metformin in a randomized double blind placebo-controlled trial. Journal of Clinical Endocrinology & Metabolism 2002;87(2):569-74
19. Garcia CR, Freeman EW, Rickels K, et al. Behavioral and emotional factors and treatment responses in a study of anovulatory infertile women. Fertility & Sterility 1985;44(4):478-83
20. Homburg R, Hendriks ML, Konig TE, et al. Clomifene citrate or low-dose FSH for the first-line treatment of infertile women with anovulation associated with polycystic ovary syndrome: a prospective randomized multinational study. Human Reproduction 2012;27(2):468-73
21. Jahan S. Comparative study of efficacy among metformin, clomiphene citrate and aromatase inhibitor (letrozole) as the first-line medication for ovulation induction, achievement of pregnancy

and live birth in Asian women with polycystic ovarian syndrome: A prospective trial. International Journal of Gynecology and Obstetrics 2015;131:E503
22. Johnson JE, Jr., Cohen MR, Goldfarb AF, et al. The efficacy of clomiphene citrate for induction of ovulation. A controlled study. Int J Fertil 1966;11(3):265-70
23. Johnson NP, Stewart AW, Falkiner J, et al. PCOSMIC: a multi-centre randomized trial in women with PolyCystic Ovary Syndrome evaluating Metformin for Infertility with Clomiphene. Human Reproduction 2010;25(7):1675-83
24. Kar S. Clomiphene citrate or letrozole as first-line ovulation induction drug in infertile PCOS women: A prospective randomized trial. Journal of human reproductive sciences 2012;5(3):262-65
25. Kar S, Sanchita S. Clomiphene citrate, metformin or a combination of both as the first line ovulation induction drug for Asian Indian women with polycystic ovarian syndrome: A randomized controlled trial. Journal of Human Reproductive Sciences 2015;8(4):197-201
26. Karimzadeh MA, Eftekhar M, Taheripanah R, et al. The effect of administration of metformin on lipid profile changes and insulin resistance in patients with polycystic ovary syndrome. Middle East Fertility Society Journal 2007;12(3):174-78
27. Karimzadeh MA, Javedani M. An assessment of lifestyle modification versus medical treatment with clomiphene citrate, metformin, and clomiphene citrate-metformin in patients with polycystic ovary syndrome. Fertility & Sterility 2010;94(1):216-20 doi: 10.1016/j.fertnstert.2009.02.078.
28. Keikha F, Shahraki Mojahed B. Induction ovulation in polycystic ovary patient with clomiphene citrate and letrozole. Iranian Journal of Reproductive Medicine 2011;9:46
29. Khorram O, Helliwell JP, Katz S, et al. Two weeks of metformin improves clomiphene citrate-induced ovulation and metabolic profiles in women with polycystic ovary syndrome. Fertility & Sterility 2006;85(5):1448-51 doi: 10.1016/j.fertnstert.2005.10.042.
30. Leanza V, Coco L, Grasso F, et al. Ovulation induction with clomiphene citrate and metformin in women with polycystic ovary syndrome. Minerva ginecologica 2014;66(3):299-301
31. Legro RS, Barnhart HX, Schlaff WD, et al. Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. New England journal of medicine 2007;356(6):551-66
32. Legro RS, Brzyski RG, Diamond MP, et al. Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. N Engl J Med 2014;371(2):119-29 doi: 10.1056/NEJMoa1313517.
33. Liu C, Feng G, Wang Q, et al. Comparison of ovulation induction protocol for women with polycystic ovarian syndrome: A prospective randomized trial. International Journal of Gynecology and Obstetrics 2015;131:E231-E32
34. Lopez E, Gunby J, Daya S, et al. Ovulation induction in women with polycystic ovary syndrome: randomized trial of clomiphene citrate versus low-dose recombinant FSH as first line therapy. Reproductive Biomedicine Online 2004;9:382-90
35. Lord J, Thomas R, Fox B, et al. The effect of metformin on fat distribution and the metabolic syndrome in women with polycystic ovary syndrome--a randomised, double-blind, placebo-controlled trial. BJOG: An International Journal of Obstetrics & Gynaecology 2006;113(7):817-24
36. Lorzadeh N, Kazemirad S, Mohammadi Z. Comparison of effects letrozole and clomiphene citrate for ovulation induction in women with polycystic ovary syndrome. Iranian Journal of Obstetrics, Gynecology and Infertility 2011;14(3)
37. Maged AM, Elsawah H, Abdelhafez A, et al. The adjuvant effect of metformin and N-acetylcysteine to clomiphene citrate in induction of ovulation in patients with Polycystic Ovary Syndrome. Gynecological endocrinology 2015
38. Mobusher I. Comparison of the efficacy of letrozole and clomiphene citrate for ovulation induction in infertile women with polycystic ovary syndrome. Pakistan Journal of Medical and Health Sciences 2014;8(4):905-08
39. Moll E, Bossuyt PM, Korevaar JC, et al. Effect of clomifene citrate plus metformin and clomifene citrate plus placebo on induction of ovulation in women with newly diagnosed polycystic ovary syndrome: randomised double blind clinical trial. BMJ (Clinical research ed) 2006;332(7556):1485 doi: 10.1136/bmj.38867.631551.55.
40. Nazik H, Kumtepe Y. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction for women with polycystic ovarian syndrome. HealthMED 2012;6(3):879-83
41. Palomba S, Orio F, Jr., Falbo A, et al. Prospective parallel randomized, double-blind, double-dummy controlled clinical trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in nonobese anovulatory women with polycystic ovary syndrome. Journal of Clinical Endocrinology & Metabolism 2005;90(7):4068-74 doi: 10.1210/jc.2005-0110.
42. Raja A, Hashmi SN, Sultana N, et al. Presentation of polycystic ovary syndrome and its management with clomiphene alone and in combination with metformin. J Ayub Med Coll

Abbottabad 2005;17(2):50-3
43. Ray PB, Ray A, Chakraborti PS. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction in Indian women with polycystic ovarian syndrome. Archives of Gynecology & Obstetrics 2012;285(3):873-7 doi: http://dx.doi.org/10.1007/s00404-011-2091-7.
44. Robinson R, Swezey M, Propst A, et al. Metformin added to clomiphene citrate does not improve pregnancy rates in hyperandrogenic, chronic anovulatory women: A randomized trial. Fertility and sterility 2003;80(Suppl 3):S273-S74, Abstract no: P
45. Roy K, Baruah J, Singla S, et al. A prospective randomized trial comparing the efficacy of Letrozole and Clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. Journal of human reproductive sciences 2012;5(1):20-25
46. Sahin Y, Yirmibes U, Kelestimur F, et al. The effects of metformin on insulin resistance, clomiphene-induced ovulation and pregnancy rates in women with polycystic ovary syndrome. European Journal of Obstetrics, Gynecology, & Reproductive Biology 2004;113(2):214-20
47. Santonocito V, Rapisarda V, Abruzzo SRM, et al. Comparison between clomiphene citrate and metformin for induction of ovulatory cycles in infertile nonobese women with polycystic ovary syndrome. Giornale Italiano di Ostetricia e Ginecologia 2009;31(11-12):455-60
48. Selim MF, Borg TF. Letrozole and clomiphene citrate effect on endometrial and subendometrial vascularity in treating infertility in women with polycystic ovary syndrome. 2012; 28(6).
49. Seyedoshohadaei F, Zandvakily F, Shahgeibi S. Comparison of the effectiveness of clomiphene citrate, tamoxifen and letrozole in ovulation induction in infertility due to isolated unovulation. Iranian Journal of Reproductive Medicine 2012;10(6):531-36
50. Sharief M, Nafee NR. Comparison of letrazole and clomiphene citrate in women with polycystic ovaries undergoing ovarian stimulation. Journal of the Pakistan Medical Association 2015;65(11):1149-52
51. Sheikh-El-Arab Elsedeek M, Elmaghraby HAH. Predictors and characteristics of letrozole induced ovulation in comparison with clomiphene induced ovulation in anovulatory PCOS women. Middle East Fertility Society Journal 2011;16(2):125-30
52. Tang T, Glanville J, Hayden CJ, et al. Combined lifestyle modification and metformin in obese patients with polycystic ovary syndrome. A randomized, placebo-controlled, double-blind multicentre study. Human Reproduction 2006;21(1):80-89
53. Vegetti W, Riccaboni A, Columbo M, et al. Randomized study of induction of ovulation by two different molecules with antioestrogenic effects, in patients with chronic anovulation disorders. 1999; 72(3 Suppl 1).
54. Williams CD, Pastore LM, Shelly WB, et al. A randomized, placebo-controlled study of the influence of instant-release metformin on response to clomiphene citrate and time to conception in polycystic ovary syndrome. Fertility and sterility 2009;92(3):S105
55. Zain MM, Jamaluddin R, Ibrahim A, et al. Comparison of clomiphene citrate, metformin, or the combination of both for first-line ovulation induction, achievement of pregnancy, and live birth in Asian women with polycystic ovary syndrome: a randomized controlled trial. Fertility & Sterility 2009;91(2):514-21
56. Zeinalzadeh M, Basirat Z, Esmailpour M. Efficacy of letrozole in ovulation induction compared to that of clomiphene citrate in patients with polycystic ovarian syndrome. Journal of reproductive medicine 2010;55(1-2):36-40

Appendix 4 List of excluded studies

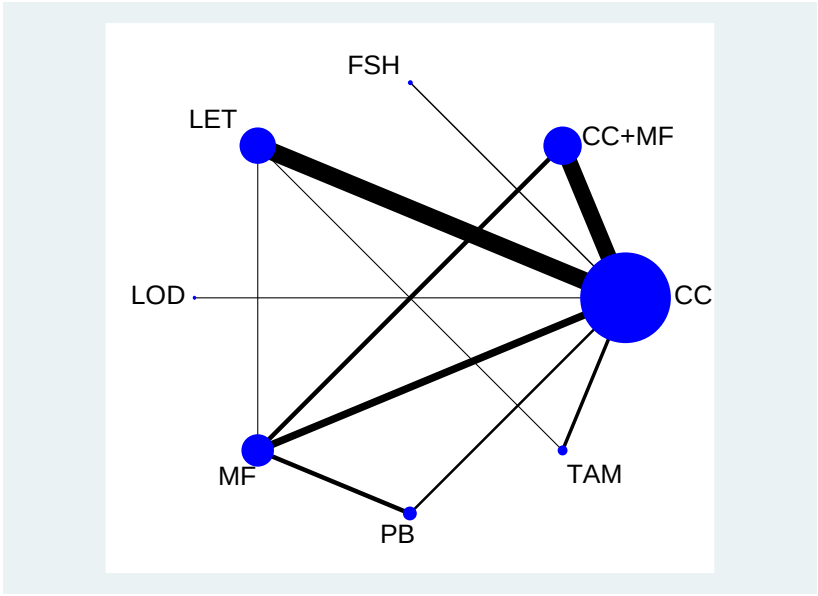
1. Aboul Enien WM, Barghash NA, Mohamed Ali FS. Clinical, ultrasonographic and endocrine predictors of ovarian response to clomiphene citrate in normogonadotropic anovulatory infertility. Middle East Fertility Society Journal 2004;9(3):242-50
2. Alamolhoda S, Mirabi P. Metformin and/or Clomiphene do not adversely affect liver or renal function in women with polycystic ovary syndrome. Iranian Journal of Reproductive Medicine 2013;11:38
3. Al-Dahhan F. The role of Metformin in induction of ovulation in obese infertile patients with polycystic ovary syndrome. International Journal of Gynecology and Obstetrics 2009;107(Journal Article):S636
4. Amer SAK, Gopalan V, Li TC, et al. Long term follow-up of patients with polycystic ovarian syndrome after laparoscopic ovarian drilling: Clinical outcome. Human Reproduction 2002;17(8):2035-42
5. Ayaz A, Alwan Y, Farooq MU. Efficacy of combined metformin-clomiphene citrate in comparison with clomiphene citrate alone in infertile women with polycystic ovarian syndrome (PCOS). Journal of Medicine & Life 2013;6(2):199-201
6. Baillargeon JP, Jakubowicz DJ, Iuorno MJ, et al. Effects of metformin and rosiglitazone, alone and in combination, in nonobese women with polycystic ovary syndrome and normal indices of insulin sensitivity. Fertil Steril 2004;82(4):893-902 doi: 10.1016/j.fertnstert.2004.02.127.
7. Baruah J, Roy KK, Rahman SM, et al. Endometrial effects of letrozole and clomiphene citrate in women with polycystic ovary syndrome using spiral artery Doppler. Archives of gynecology and obstetrics 2009;279(3):311-4 doi: 10.1007/s00404-008-0714-4.
8. Basirat Z, Golsourkhtabar M, Kashifard M. Does metformin affect on pregnancy outcome in PCOs infertile women with different BMI? Iranian Journal of Reproductive Medicine 2012;10:76
9. Begum MR, Ferdous J, Begum A, et al. Comparison of efficacy of aromatase inhibitor and clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. Fertility and sterility 2009;92(3):853-57 doi: 10.1016/j.fertnstert.2007.08.044 [doi].
10. Beigi A, Zarrinkoub F. Randomized controlled trial comparing clomiphene citrate and metformin as the first-line treatment for ovulation induction in women with polycystic ovary syndrome. XVIII FIGO World Congress of Gynecology and Obstetrics 2006;3:177
11. Ben Ayed B, Dammak dit Mlik S, Ben Arab H, et al. Metformin effects on clomifene-induced ovulation in the polycystic ovary syndrome. Tunisie Medicale 2009;87(1):43-9
12. Connaughton JF, Jr., Garcia CR, Wallach EE. Induction of ovulation with cisclomiphene and a placebo. Obstetrics & Gynecology 1974;43(5):697-701
13. Diamond MP, Kruger M, Santoro N, et al. Adverse impact of progestin exposure and endometrial shedding prior to ovulation induction on conception and live birth in women with polycystic ovary syndrome. Reproductive Sciences 2012;19(3):97A
14. El Bigawy AF, Fouda UMF, Wahab HAE. A randomized trial of letrazole versus clomiphene citrate in induction of ovulation in patients with polycystic ovary syndrome (PCOS). Middle East Fertility Society Journal 2008;13(1):52-56
15. Ghahiri A, Mamorian M. Comparative study of aromatase inhibitor (Letrozole) with clomiphen citrate as the first line treatment of patients with PCO. Iranian Journal of Reproductive Medicine 2010;8:84
16. Hashim HA, Bazeed M, Elaal IA. Minimal stimulation or clomiphene citrate as first-line therapy in women with polycystic ovary syndrome: A randomized controlled trial. Gynecological endocrinology 2012;28(2):87-90
17. Homburg R, Hendriks ML, Konig T, et al. Clomifene or low-dose FSH for the first-line treatment of anovulatory PCOS: A prospective randomised multinational study (COFFI). Molecular Human Reproduction 2009;24:i22-i23
18. Hosseini MA, Alleyassin A, Sarvi F, et al. Metformin treatment in different phenotypes of polycystic ovary syndrome. Archives of Gynecology and Obstetrics 2013;288(5):1131-36
19. Imani B, Eijkemans MJC, Te Velde ER, et al. A nomogram to predict the probability of live birth after clomiphene citrate induction of ovulation in normogonadotropic oligoamenorrheic infertility. Fertility and sterility 2002;77(1):91-97
20. Johnson N. PCOSMIC polycystic ovarian syndrome, metformin for infertility with clomiphene: a muli centre randomised controlled trial. FSA Fertililty Society Australia, Abstract Book of Proceedings Brisbane, Australia 19-22 October 2008 2008:34
21. Johnson NP. PCOSMIC-polycystic ovarian syndrome, metformin for infertility with clomiphene: A multi-centre double-blind randmised controlled trial. Molecular Human Reproduction 2009;24:i24

22. Jungheim ES, Odibo AO. Fertility treatment in women with polycystic ovary syndrome: A decision analysis of different oral ovulation induction agents. Fertility and sterility 2010;94(7):2659-64
23. Kar S. Clomiphene citrate or letrozole as first line ovulation induction drug in infertile PCOS women: A prospective randomised trial. Fertility and sterility 2012;98(3):S86
24. Kar S. Clomiphene citrate, metformin or the combination of both, as first line ovulation induction drug in polycystic ovarian syndrome: A randomised controlled trial. Fertility and sterility 2013;100(3):S359-S60
25. Ladson G, Dodson WC, Sweet SD, et al. The effects of metformin with lifestyle therapy in polycystic ovary syndrome: a randomized double-blind study. Fertil Steril 2011;95(3):1059-66 e1 doi: 10.1016/j.fertnstert.2010.12.002.
26. Legro RS, Brzyski RG, Diamond MP, et al. The Pregnancy in Polycystic Ovary Syndrome II study: baseline characteristics and effects of obesity from a multicenter randomized clinical trial. Fertility & Sterility 2014;101(1):258-69.e8 doi: http://dx.doi.org/10.1016/j.fertnstert.2013.08.056.
27. Legro RS, Brzyski RG, Diamond MP, et al. Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. Obstetrical and Gynecological Survey 2015;69(10):599-601
28. Legro RS, Kunselman AR, Brzyski RG, et al. The Pregnancy in Polycystic Ovary Syndrome II (PPCOS II) trial: rationale and design of a double-blind randomized trial of clomiphene citrate and letrozole for the treatment of infertility in women with polycystic ovary syndrome. Contemporary clinical trials 2012;33(3):470-81 doi: 10.1016/j.cct.2011.12.005.
29. Legro RS, Myers ER, Barnhart HX, et al. The Pregnancy in Polycystic Ovary Syndrome Study: baseline characteristics of the randomized cohort including racial effects. Fertility and Sterility 2006;86(4):914-33 doi: http://dx.doi.org/10.1016/j.fertnstert.2006.03.037.
30. Lorzadeh N, Kazemirad S. Comparison of effects letrozole and clomiphene citrate for ovulation induction in women with polycystic ovary syndrome. Journal fur Reproduktionsmedizin und Endokrinologie 2010;7(4):367
31. Ma LK, Jin LN, Yu Q, et al. [Effect of lifestyle adjustment, metformin and rosiglitazone in polycystic ovary syndrome]. Zhonghua Fu Chan Ke Za Zhi 2007;42(5):294-7
32. Moll E, Korevaar JC, Bossuyt PMM, et al. Does adding metformin to clomifene citrate lead to higher pregnancy rates in a subset of women with polycystic ovary syndrome? Human Reproduction 2008;23(8):1830-34 doi: 10.1093/humrep/den182.
33. Moll E, van Wely M, Lambalk CB, et al. Health-related quality of life in women with newly diagnosed polycystic ovary syndrome randomized between clomifene citrate plus metformin or clomifene citrate plus placebo. Human reproduction (Oxford, England) 2012;27(11):3273-8 doi: 10.1093/humrep/des310.
34. Morin-Papunen L, Rantala AS, Unkila-Kallio L, et al. Metformin improves pregnancy and live-birth rates in women with polycystic ovary syndrome (PCOS): a multicenter, double-blind, placebo-controlled randomized trial. Journal of Clinical Endocrinology & Metabolism 2012;97(5):1492-500
35. Moslemizadeh N, Moghadam TG, Ehteshami S. Comparison of clomiphene citrate plus estradiol, with tamoxifen citrate effects in induction of ovulation and pregnancy in poly cystic ovarian syndrome patients. Journal of Medical Sciences 2008;8(8):734-38
36. Nahid L, Sirous K. Comparison of the effects of letrozole and clomiphene citrate for ovulation induction in infertile women with polycystic ovary syndrome. [Italian]. Minerva ginecologica 2012;64(3):253-8
37. NIH/NICHD Reproductive Medicine Network. Letrozole versus clomiphene citrate in a novulatory PCOS women: A cost-effectiveness analysis. Fertility and sterility 2013;100(3):S128
38. NIH/NICHDReproductive Medicine Network. Effect of Letrozole versus Clomiphene on live birth in women with anovulatory infertility due to polycystic ovary syndrome (PCOS): a randomized double-blind multi-center trial. Fertility and sterility 2013;100 Suppl(3):S51
39. Rausch ME, Legro RS, Barnhart HX, et al. Predictors of pregnancy in women with polycystic ovary syndrome. Journal of Clinical Endocrinology & Metabolism 2009;94(9):3458-66 doi: 10.1210/jc.2009-0545.
40. Ridzuan J, Murizah MZ, Adibah I, et al. Prospective randomized control trial comparing metformin and clomiphine citrate as ovulation induction agent in women with polycystic ovarian syndrome at Alor Star hospital, Malaysia. XVIII FIGO World Congress of Gynecology and Obstetrics 2006;3:181-2
41. Schlaff D, Legro RS, Diamond MP, et al. Ovulation induction with clomiphene, metformin, or acombination of the two does not affect hirsutism score over a standard course of treatment in women with polycystic ovary syndrome. Fertility and sterility 2010;94(4):S192-S93

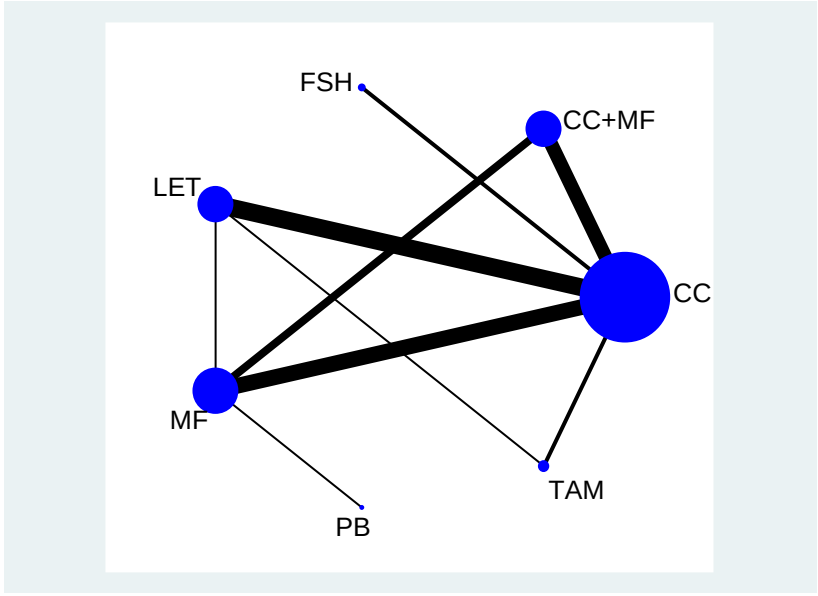
42. Seyedoshohadaei F, Zandvakily F, Shahgeibi SH. Effectiveness of clomiphene citrate, tamoxifen and letrozole in induction ovulation in women with infertillity due to isolated unovulation. Iranian Journal of Reproductive Medicine 2011;9:33
43. Siebert TI, Kruger TF, Lombard C. Evaluating the equivalence of clomiphene citrate with and without metformin in ovulation induction in PCOS patients. Journal of Assisted Reproduction & Genetics 2009;26(4):165-71
44. Topc¸u HO, Yüksel B, Islimye M, et al. Tamoxifen as a first line treatment in ovulation induction for the infertile women with thin endometrium and high follicular response with clomiphene. Human Reproduction 2010;25:i312
45. Wang L, Lu S, Cao Z, et al. Effects of letrozole and clomiphene citrate on ovulation induction in polycystic ovarian syndrome. Journal of Xi'an Jiaotong University (Medical Sciences) 2008;29(6):69

Appendix 5 (a-e) Network plots of eligible comparisons for five outcomes: pregnancy, live birth, ovulation, miscarriage and multiple pregnancy.
The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of studies including the respective interventions. (Abbreviations: CC, clomiphene citrate; PB, placebo or no treatment; LET, letrozole; MF, metformin; TAM, tamoxifen; FSH, follicle stimulating hormone; LOD, laparoscopic ovarian drilling)

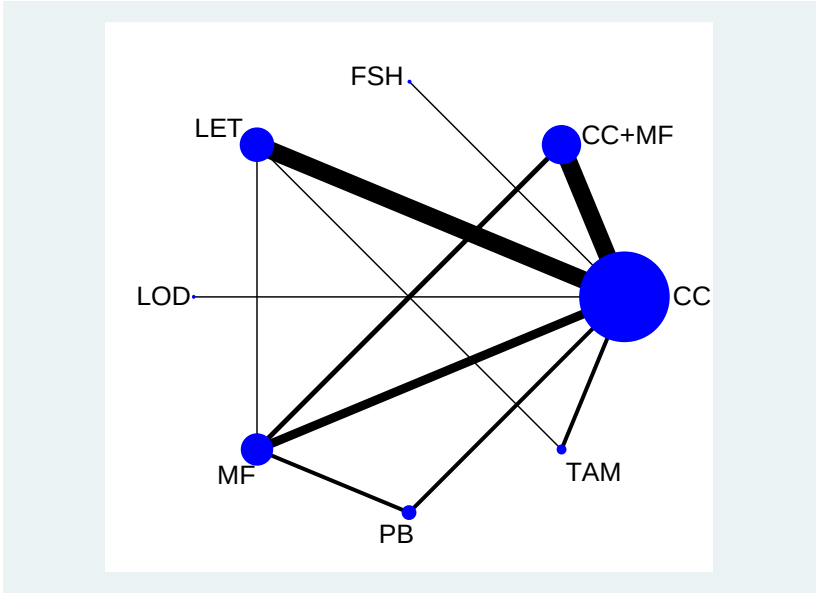
5a. Pregnancy



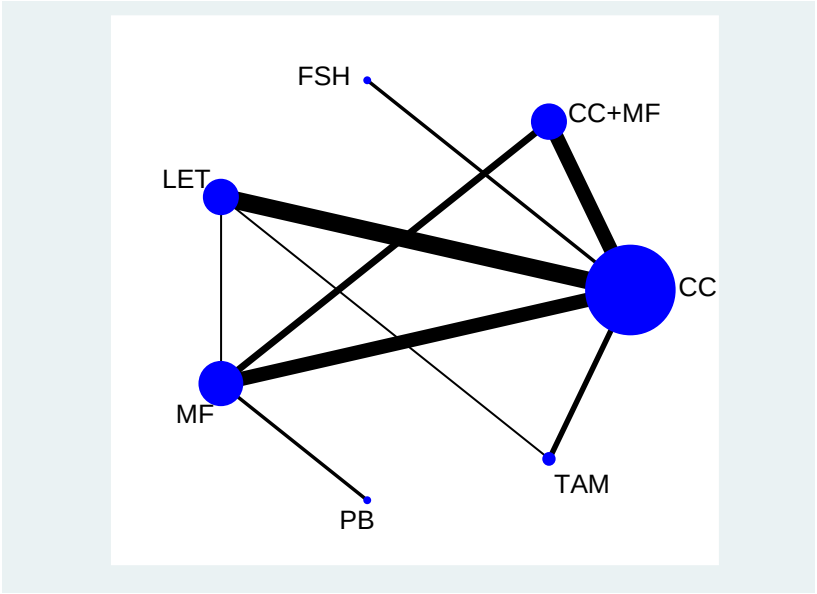
5b. live birth



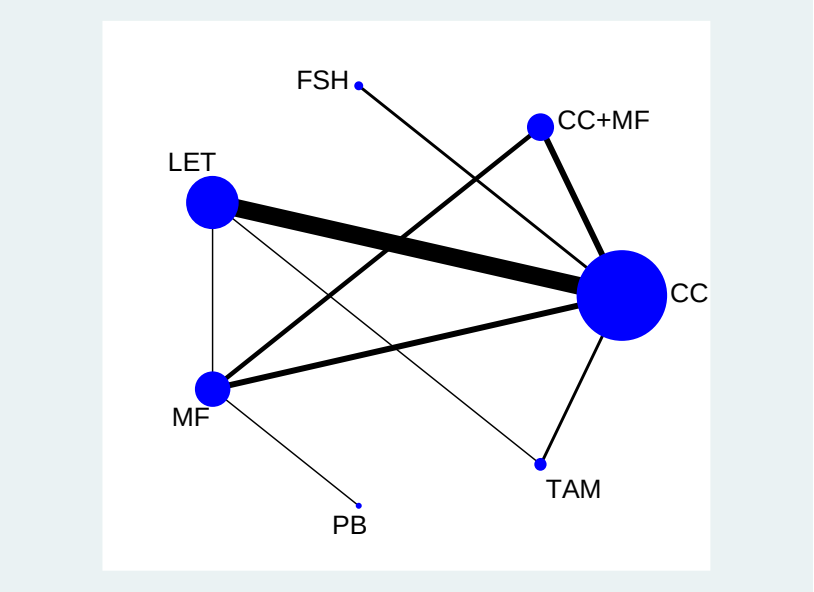
5c. ovulation



5d. miscarriage

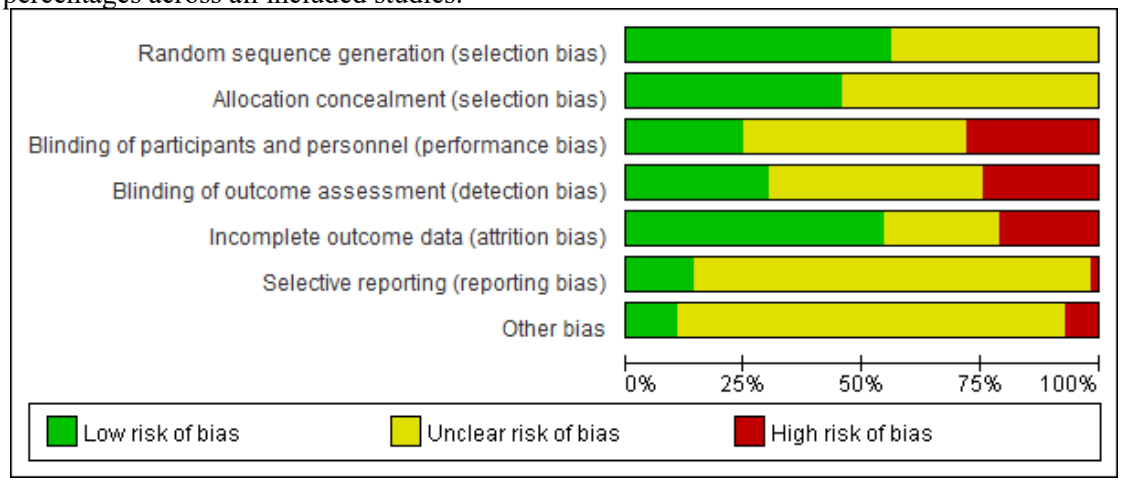


5e. multiple pregnancy

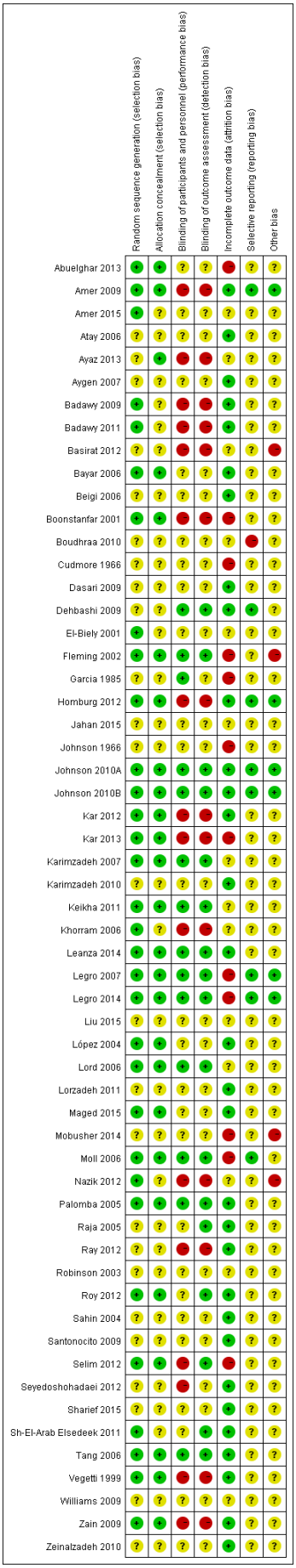


Appendix 6 Risk of bias evaluation.

6a. Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies.



6b. Risk of bias summary: review authors' judgments about each risk of bias item for each included study.


Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Incomplete outcome data (attrition bias)
Selective reporting (reporting bias)
Other bias
Abuelghar 2013
Amer 2009
Amer 2015
Atay 2006
Ayaz 2013
Aygen 2007
Badawy 2009
Badawy 2011
Basirat 2012
Bayar 2006
Beigi 2006
Boonstanfar 2001
Boudhraa 2010
Cudmore 1966
Dasari 2009
Dehbashi 2009
El-Biely 2001
Fleming 2002
Garcia 1985
Homburg 2012
Jahan 2015
Johnson 1966
Johnson 2010A
Johnson 2010B
Kar 2012
Kar 2013
Karimzadeh 2007
Karimzadeh 2010
Keikha 2011
Khorram 2006
Leanza 2014
Legro 2007
Legro 2014
Liu 2015
López 2004
Lord 2006
Lorzadeh 2011
Maged 2015
Mobusher 2014
Moll 2006
Nazik 2012
Palomba 2005
Raja 2005
Ray 2012
Robinson 2003
Roy 2012
Sahin 2004
Santonocito 2009
Selim 2012
Seyedoshohadaei 2012
Sharief 2015
Sh-El-Arab Elsedeek 2011
Tang 2006
Vegetti 1999
Williams 2009
Zain 2009
Zeinalzadeh 2010

Appendix 7 Pairwise meta-analysis results for direct comparisons of interventions

Comparisons	Pairwise meta-analysis odds ratio (95% CI)	No. of trials	No. of participants	Heterogeneity I^2
Pregnancy				
PB vs CC	0.20(0.05-0.74)	3	136	0%
LET	1.52(1.26-1.85)	21	3553	24.3%
MF	1.10(0.62-1.95)	9	1335	73.1%
CC+MF	1.56(1.24-1.97)	19	2070	12.2%
TAM	0.64(0.36-1.12)	4	661	43.7%
FSH	1.57(1.04-2.37)	2	378	0%
LOD	0.52(0.19-1.44)	1	72	N/A
MF vs PB	3.58(2.06-6.21)	5	494	0%
MF vs LET	0.73(0.41-1.32)	1	304	N/A
TAM	0.67(0.30-1.47)	1	100	N/A
CC+MF vs MF	1.92(0.90-4.06)	5	818	71.8%
Live Birth				
LET vs CC	1.60(1.30-1.98)	9	1990	0%
MF	1.00(0.45-2.22)	8	1155	80.9%
CC+MF	1.14(0.81-1.61)	7	950	12.4%
TAM	0.96(0.26-3.55)	2	195	35.3%
FSH	1.50(0.98-2.29)	2	378	0%
MF vs PB	2.87(0.51-16.02)	1	65	N/A
MF vs LET	0.38(0.19-0.78)	1	304	N/A
TAM	0.71(0.32-1.60)	1	100	N/A
CC+MF vs MF	2.48(1.24-4.95)	4	640	51.1%
Ovulation (per woman randomised)				
PB vs CC	0.15(0.07-0.34)	3	136	0%
LET	1.89(1.55-2.30)	14	2568	8.8%
MF	0.62(0.32-1.22)	7	1119	82.9%
CC+MF	1.46(1.01-2.12)	14	1407	54.5%
TAM	0.61(0.43-0.86)	3	566	0%
FSH	0.11(0.76-12.79)	1	76	N/A
LOD	0.70(0.27-1.83)	1	72	N/A
MF vs PB	3.63(0.45-29.35)	3	309	92.9%
MF vs LET	0.14(0.09-0.24)	1	304	N/A
TAM	0.75(0.31-1.78)	1	100	N/A
CC+MF vs MF	3.20(1.85-5.52)	4	640	44.4%
Multiple pregnancy (per woman randomised)				
LET vs CC	0.45(0.22-0.91)	12	2460	0%
MF	0.22(0.05-0.96)	4	976	0%
CC+MF	0.57(0.19-1.74)	4	892	0%
TAM	0.48(0.06-3.76)	2	471	0%
FSH	3.62(0.58-22.80)	2	378	0%
MF vs PB	0.33(0.01-8.49)	1	65	N/A
MF vs LET	0.20(0.01-4.15)	1	304	N/A
TAM	3.06(0.12-76.95)	1	100	N/A
CC+MF vs MF	2.36(0.42-12.39)	4	665	0%
Miscarriage (per woman randomised)				
LET vs CC	1.00(0.62-1.62)	10	2302	10.6%
MF	0.76(0.32-1.82)	8	1155	29.1%
CC+MF	1.38(0.85-2.24)	8	991	0%
TAM	0.56(0.19-1.68)	3	566	23.4%
FSH	1.44(0.57-3.63)	2	378	0%
MF vs PB	1.02(0.28-3.73)	2	265	0%
MF vs LET	0.33(0.13-8.20)	1	304	N/A
TAM	0.73(0.16-3.46)	1	100	N/A

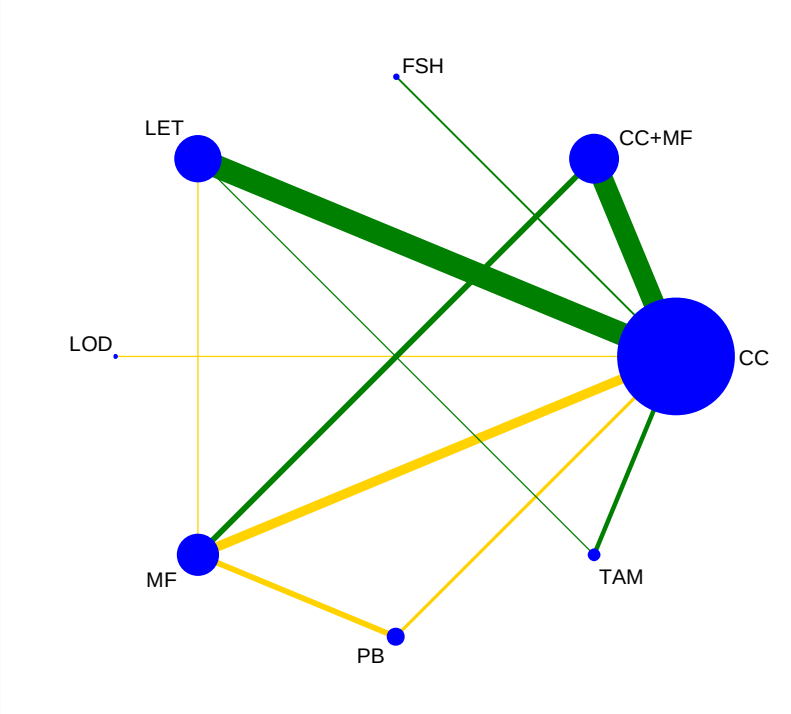
CC+MF	**vs MF**	1.37(0.66-2.87)	4	640	10.9%
Miscarriage (per pregnant woman)					
LET	**vs CC**	0.79(0.52-1.21)	10	718	0%
MF		0.70(0.19-2.63)	8	277	54.9%
CC+MF		1.35(0.74-2.46)	8	384	0%
TAM		0.83(0.31-2.19)	3	123	0%
FSH		0.99(0.37-2.67)	2	164	0%
MF	**vs PB**	0.28(0.06-1.19)	2	63	0%
MF	**vs LET**	0.41(0.02-10.64)	1	55	N/A
TAM		0.93(0.18-4.72)	1	45	N/A
CC+MF	**vs MF**	0.67(0.27-1.66)	4	174	0%

(Abbreviations: CC, clomiphene citrate; PB, placebo or no treatment; LET, letrozole; MF, metformin; TAM, tamoxifen; FSH, follicle stimulating hormone; LOD, laparoscopic ovarian drilling)

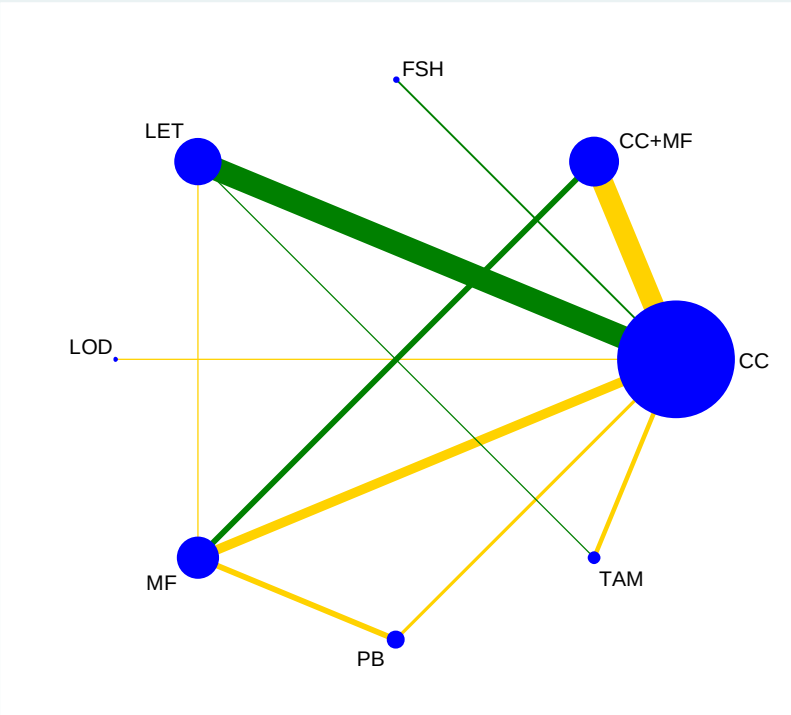
Appendix 8 Network plot for pregnancy incorporating risk of bias assessment

8a. Risk of bias in randomisation

Colored edges are based on adequacy of randomisation in the majority of the trials in each comparison. Green, yellow and red colors represent low, unclear and high risk, respectively.

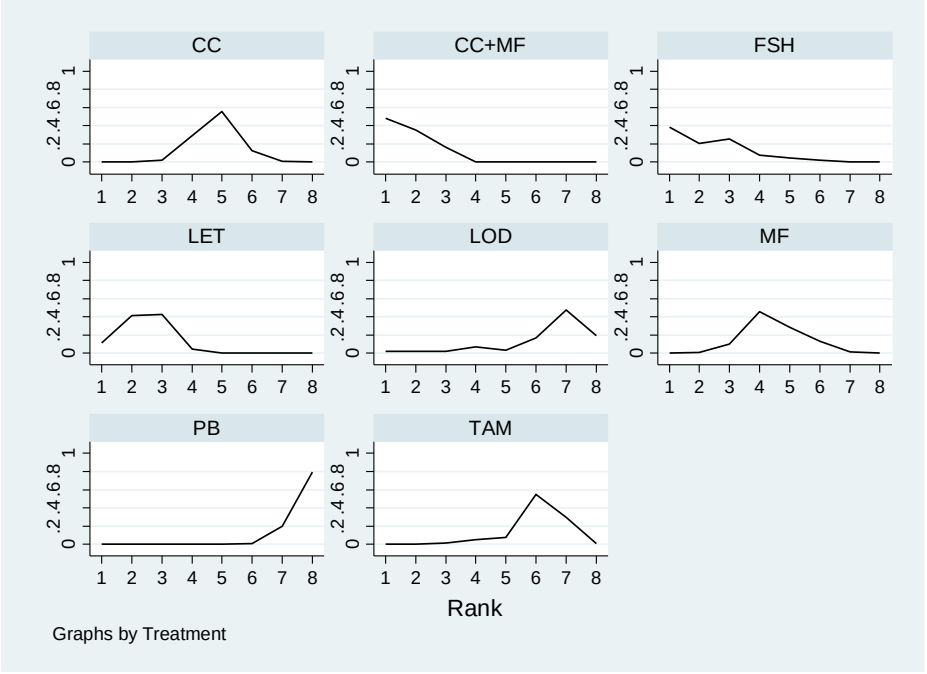


8b. Risk of randomisation in allocation concealment



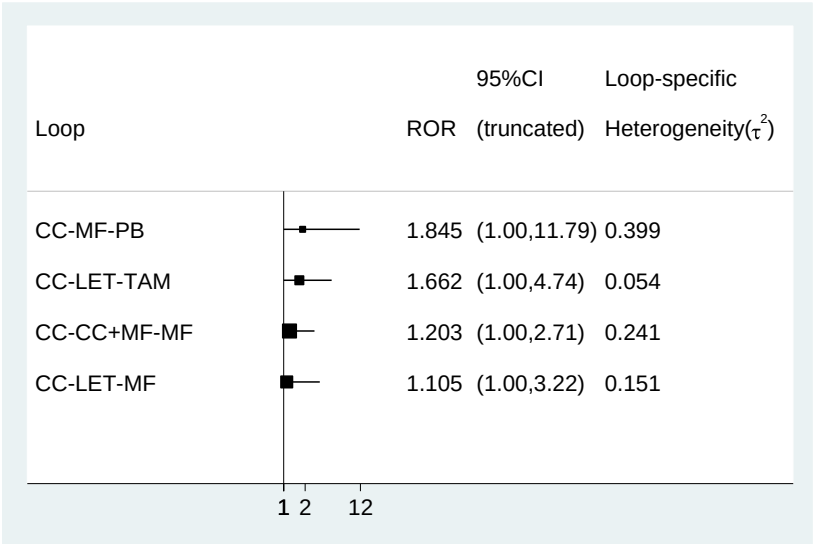
Appendix 9 Ranking of treatments for pregnancy

Rankograms below illustrate the probability per rank for each treatment in terms of pregnancy. E.g. for CC, the probabilities of being the best treatment, the second best, to the worst (eighth) are 0%, 0%, 2.4%, 29.0%, 55.5%, 12.3%, 0.8% and 0%, respectively.

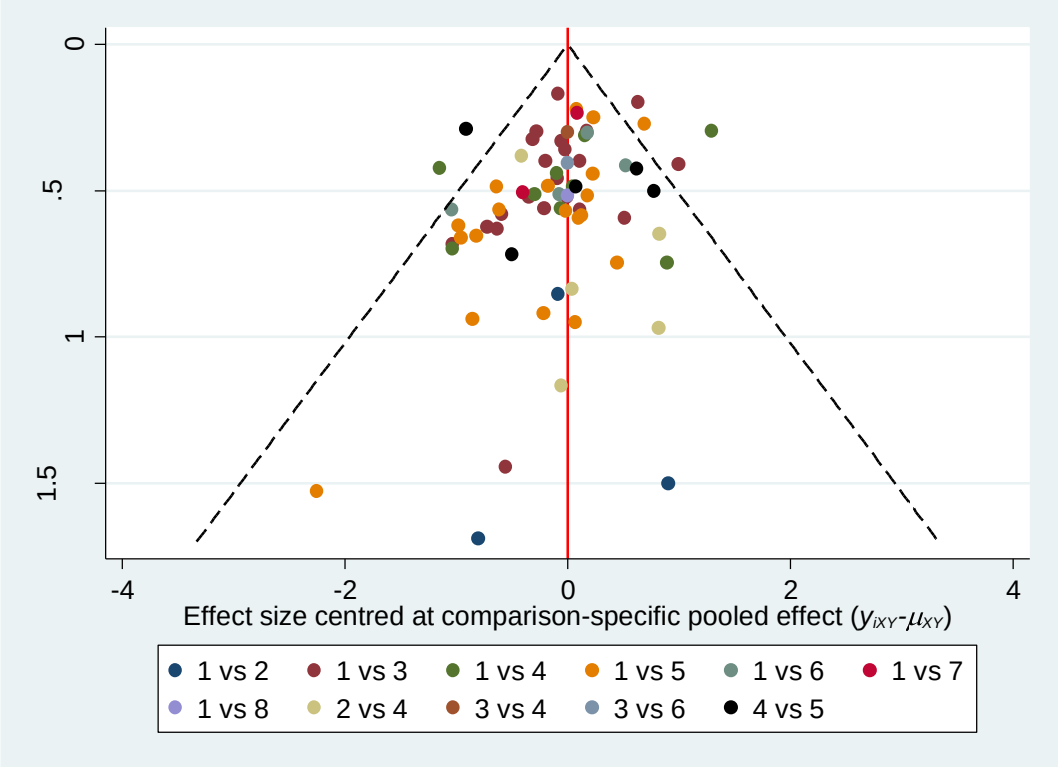


Appendix 10 Inconsistency plot for pregnancy

We estimated inconsistency as the logarithm of the ratio of two odds ratios (RoR) from direct and indirect evidence in the loop (also named inconsistency factor IF) and the corresponding 95% CI for each IF in each closed triangular or quadratic loop. RoR values is close to 1 mean that the two sources are in agreement. The inconsistency plot shows that in a total of 4 loops there is none with statistically significant inconsistency as all confidence intervals for RORs are compatible with zero inconsistency (RoR= 1).

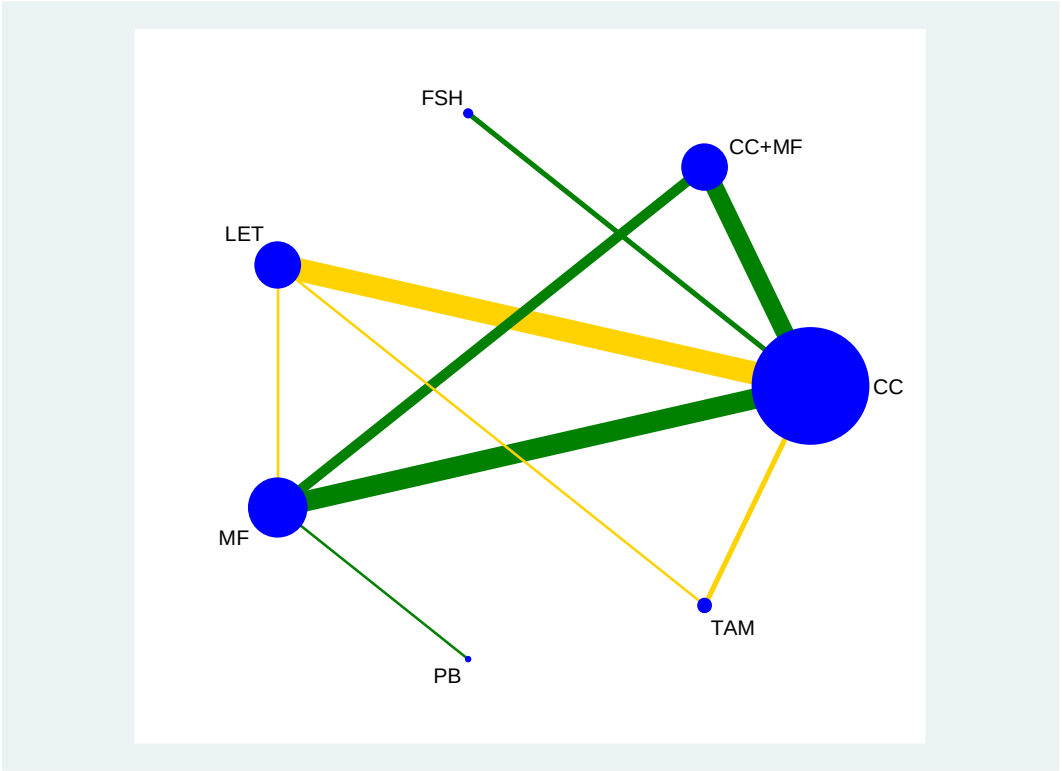


Appendix 11 Comparison-adjusted funnel plot for pregnancy
The red line represents the null hypothesis that the study-specific effect sizes do not differ from the respective comparison-specific pooled effect estimates. Different colors correspond to different comparisons. (1-clomiphene; 2-placebo/no treatment; 3-letrozole; 4-metformin; 5-clomiphene plus metformin; 6-tamoxifen; 7-FSH; 8-LOD)

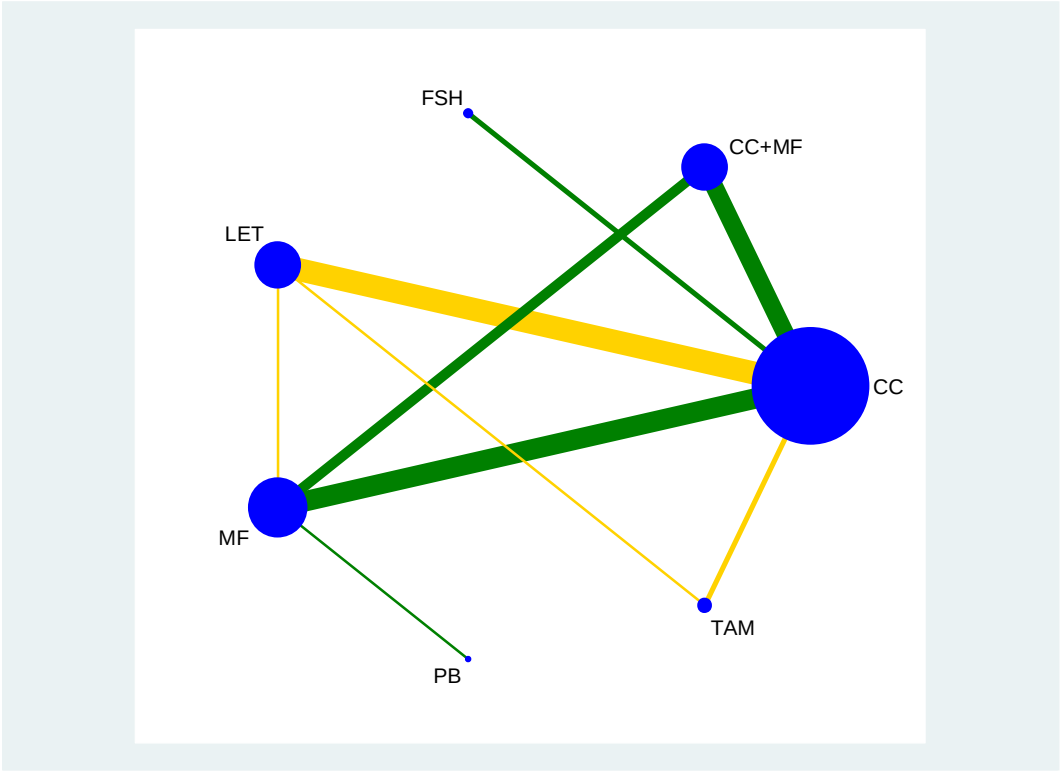


Appendix 12 Network plot for live birth incorporating risk of bias assessment

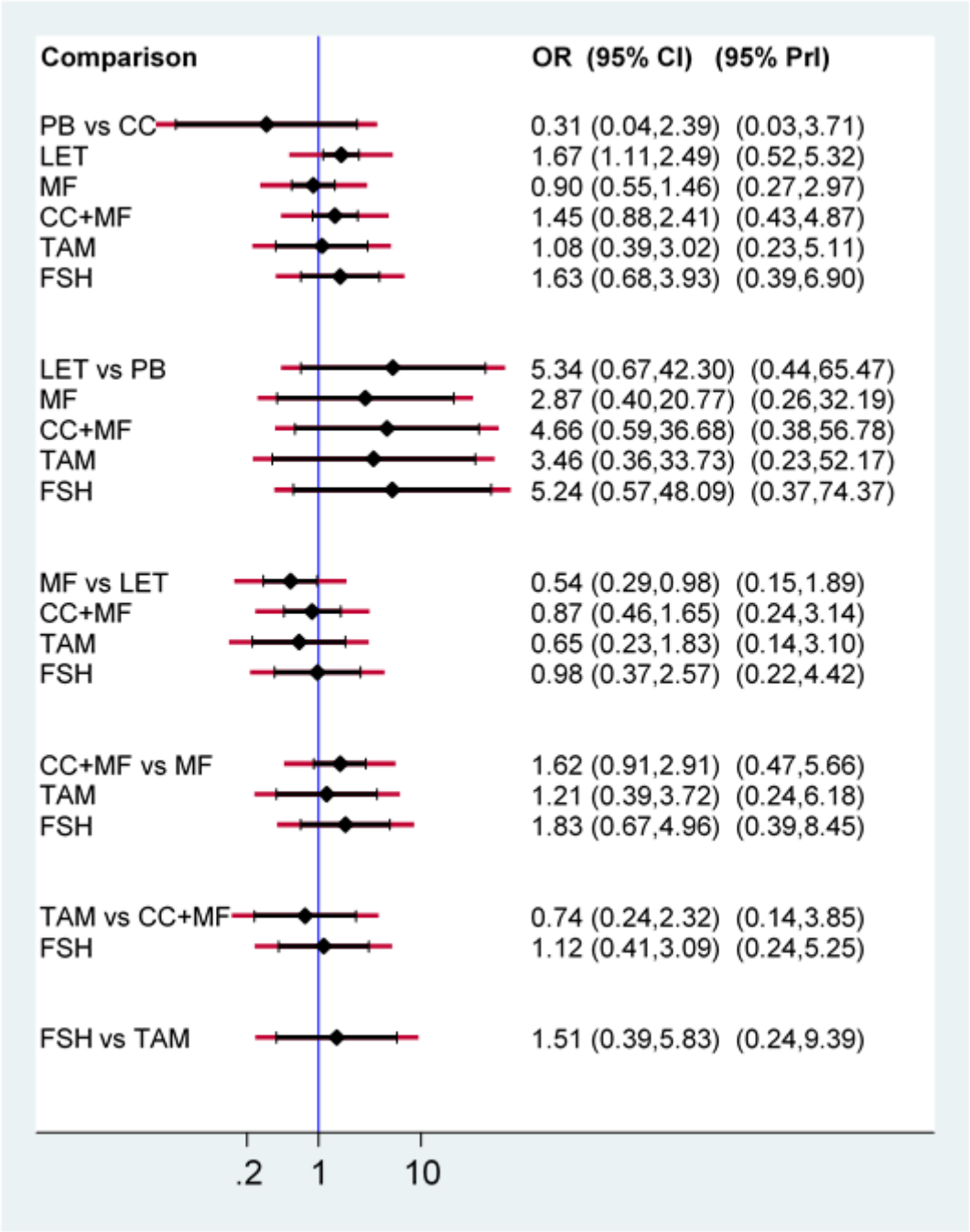
12a. Risk of bias in randomisation



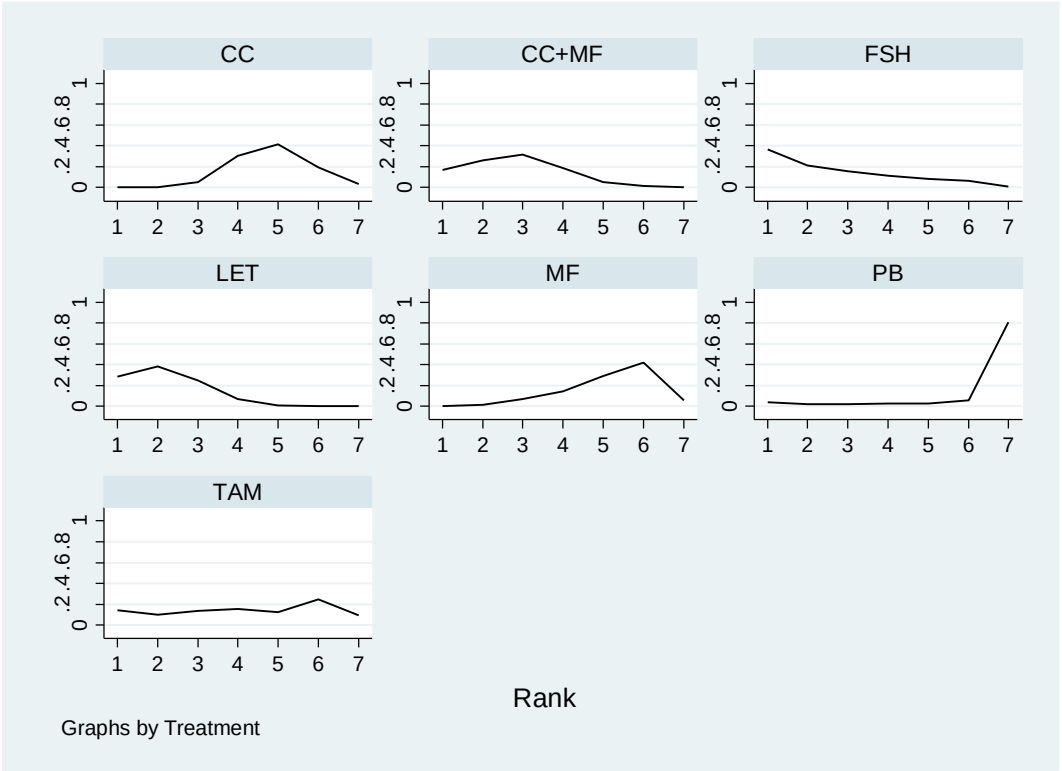
12b. Risk of bias in allocation concealment



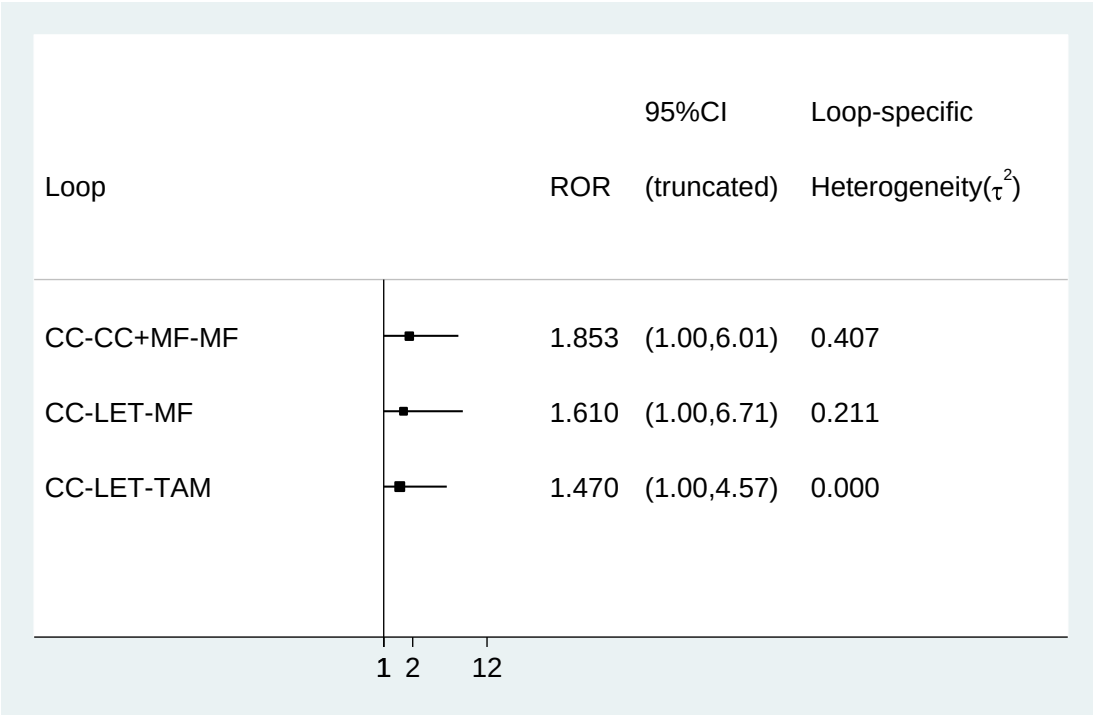
Appendix 13 Network meta-analysis results for live birth



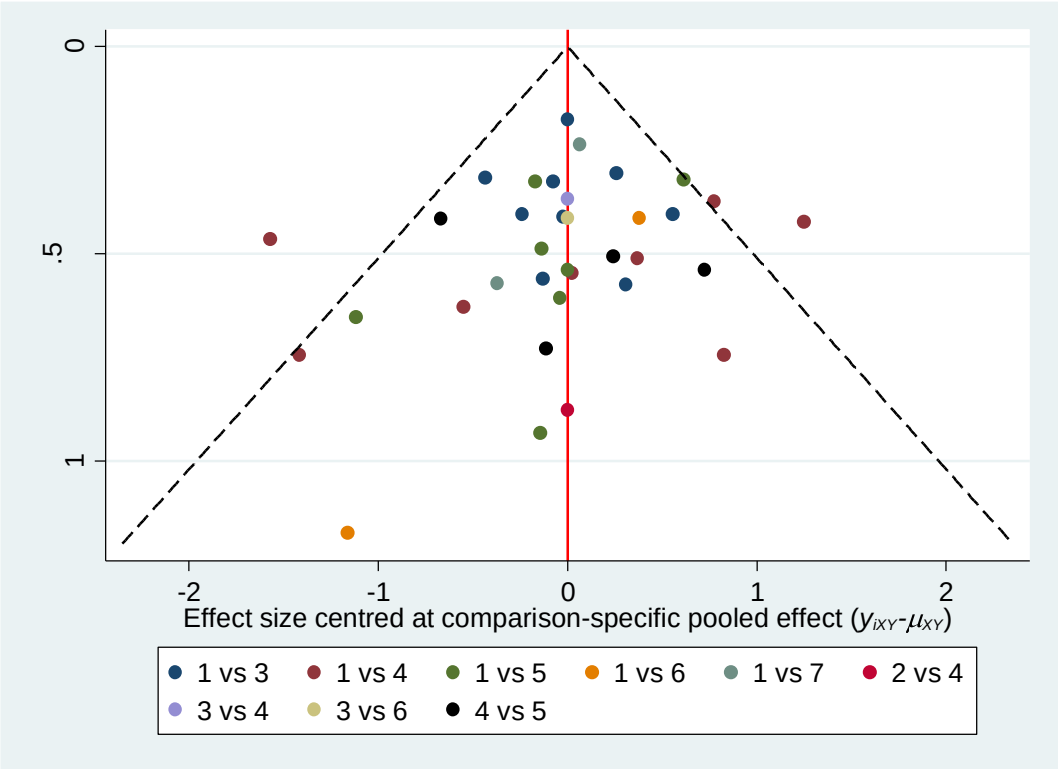
Appendix 14 Ranking of treatments for live birth



Appendix 15 Inconsistency plot for live birth.
The inconsistency plot shows that in a total of 3 loops there is none with statistically significant inconsistency as all confidence intervals for RORs are compatible with zero inconsistency (RoR= 1).



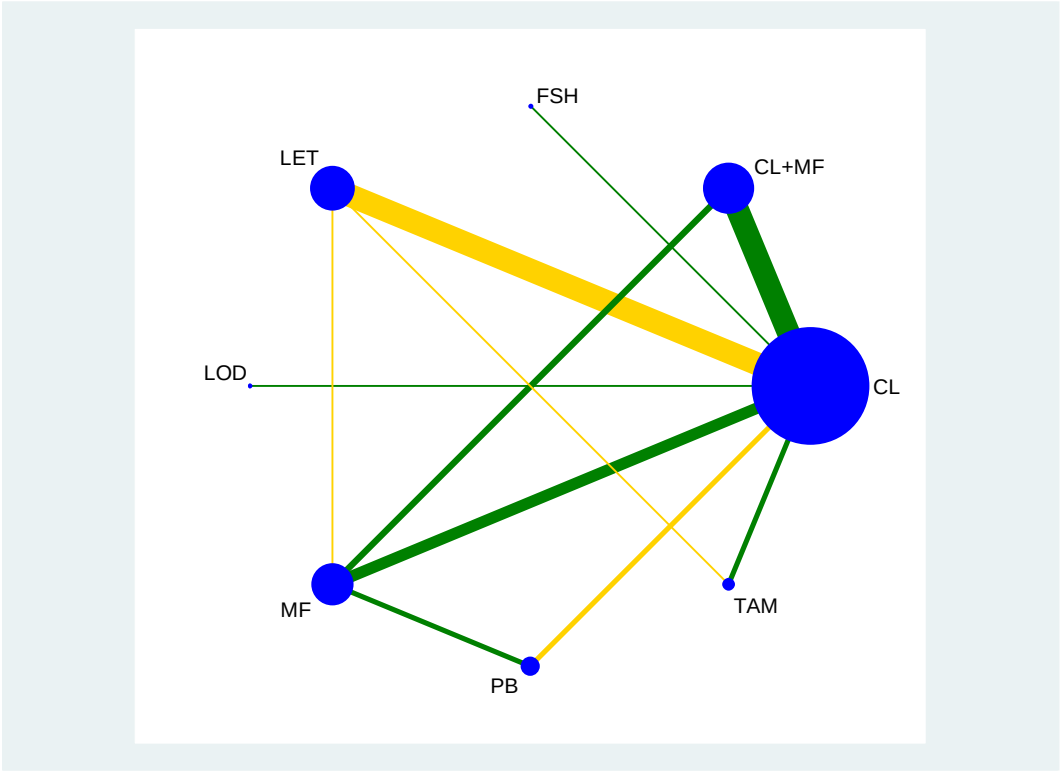
Appendix 16 Comparison-adjusted funnel plot for live birth



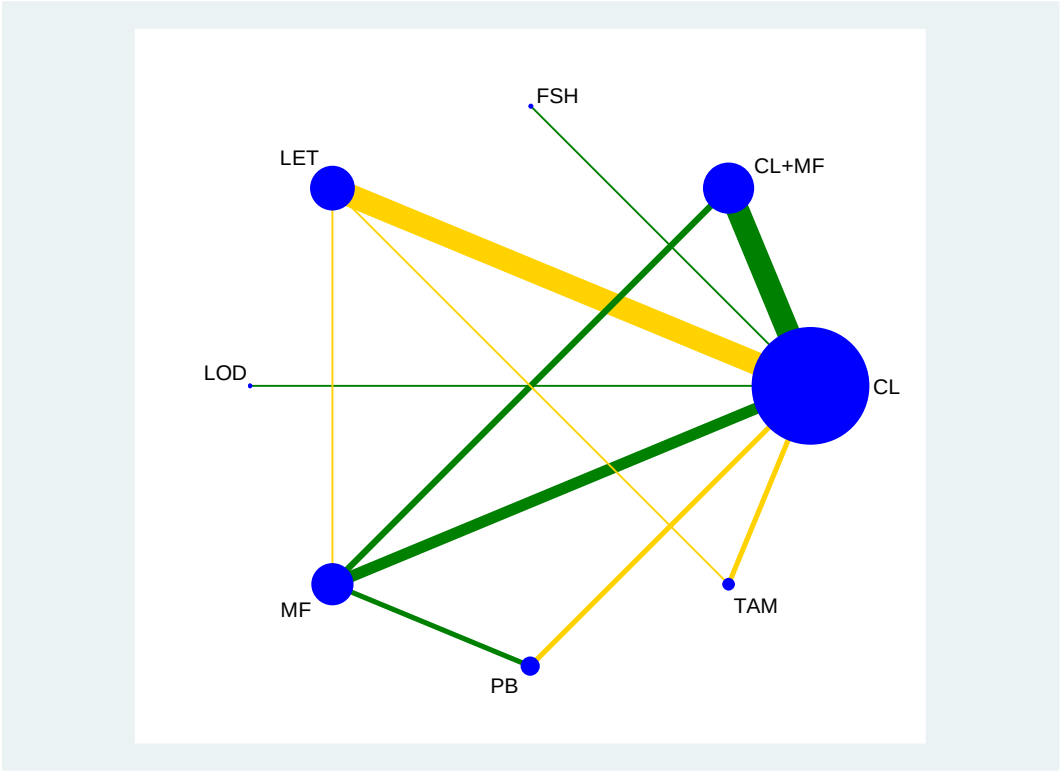
(1-clomiphene; 2-placebo/no treatment; 3-letrozole; 4-metformin; 5-clomiphene plus metformin; 6-tamoxifen; 7-FSH)

Appendix 17 Network plot for ovulation incorporating risk of bias assessment

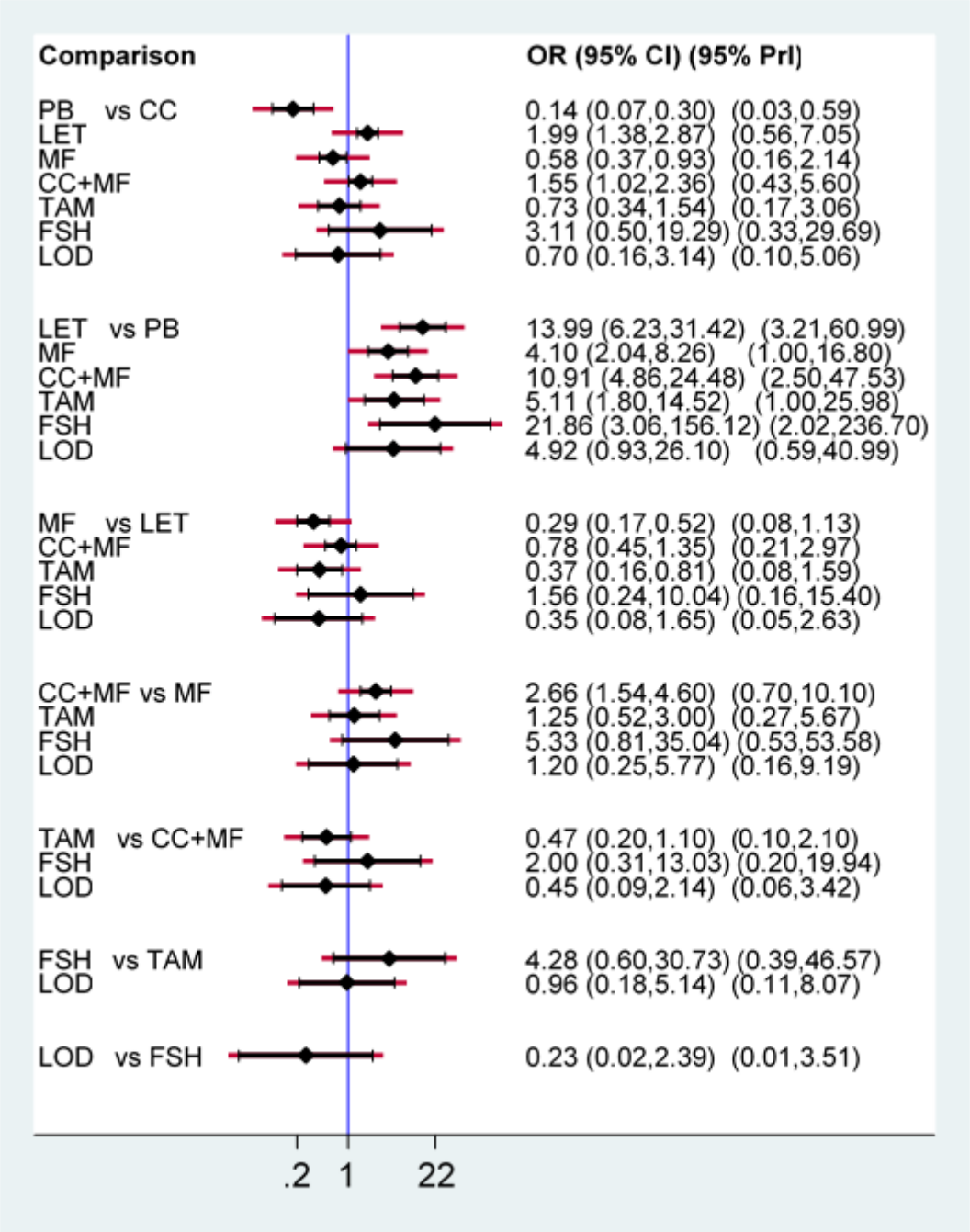
17a. Risk of bias in randomisation



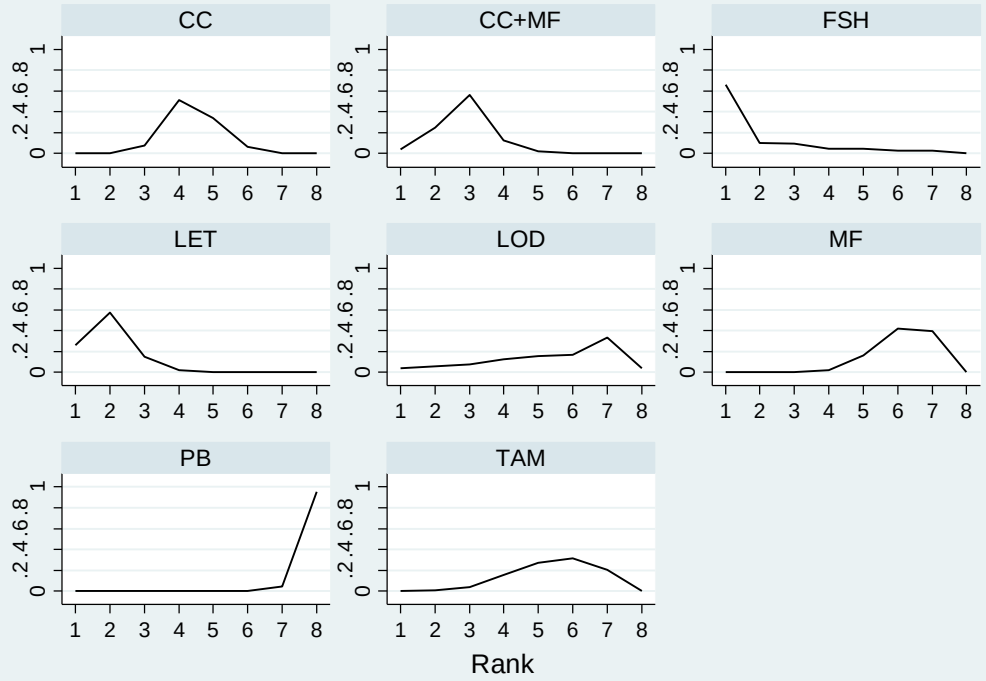
17b. Risk of bias in allocation concealment



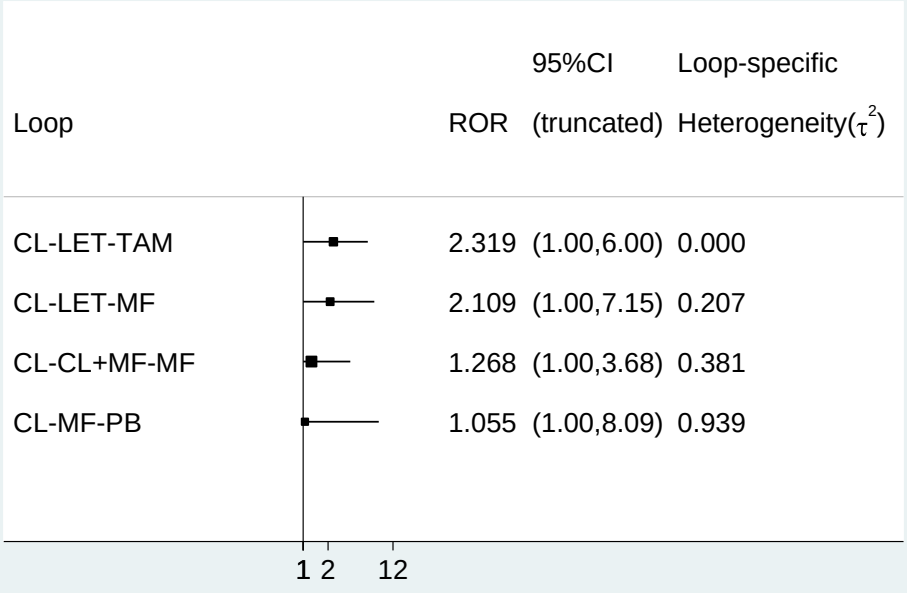
Appendix 18 Network meta-analysis results for ovulation



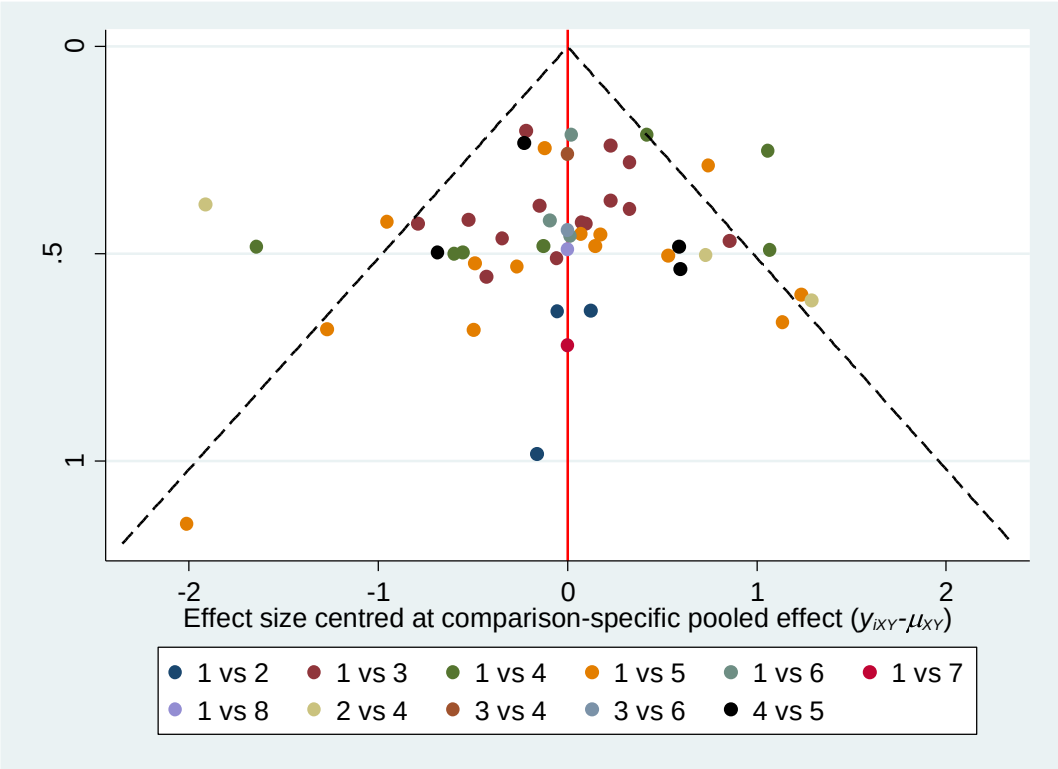
Appendix 19 Ranking of treatments for ovulation



Appendix 20 Inconsistency plot for ovulation



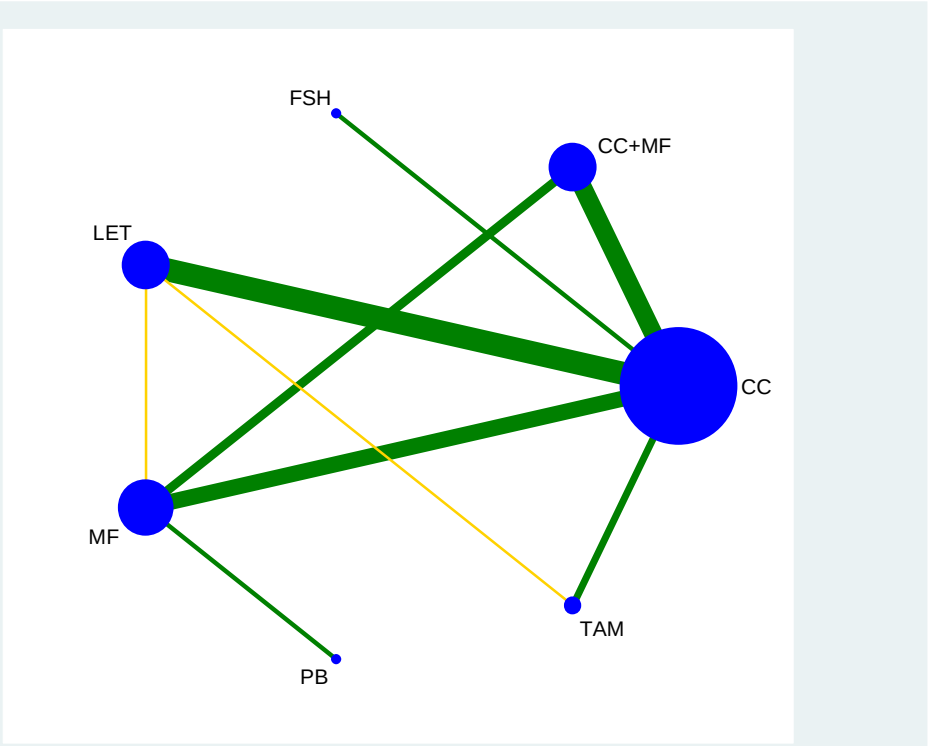
Appendix 21 Comparison-adjusted funnel plot for ovulation



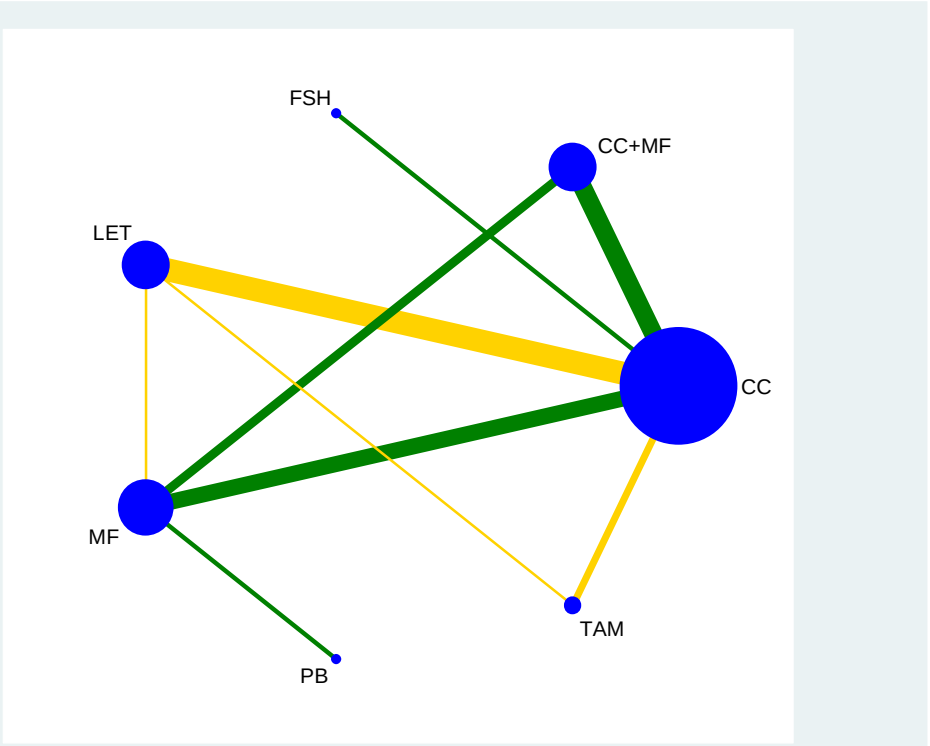
(1-clomiphene; 2-placebo/no treatment; 3-letrozole; 4-metformin; 5-clomiphene plus metformin; 6-tamoxifen; 7-FSH; 8-LOD)

Appendix 22 Network plot for miscarriage incorporating risk of bias assessment

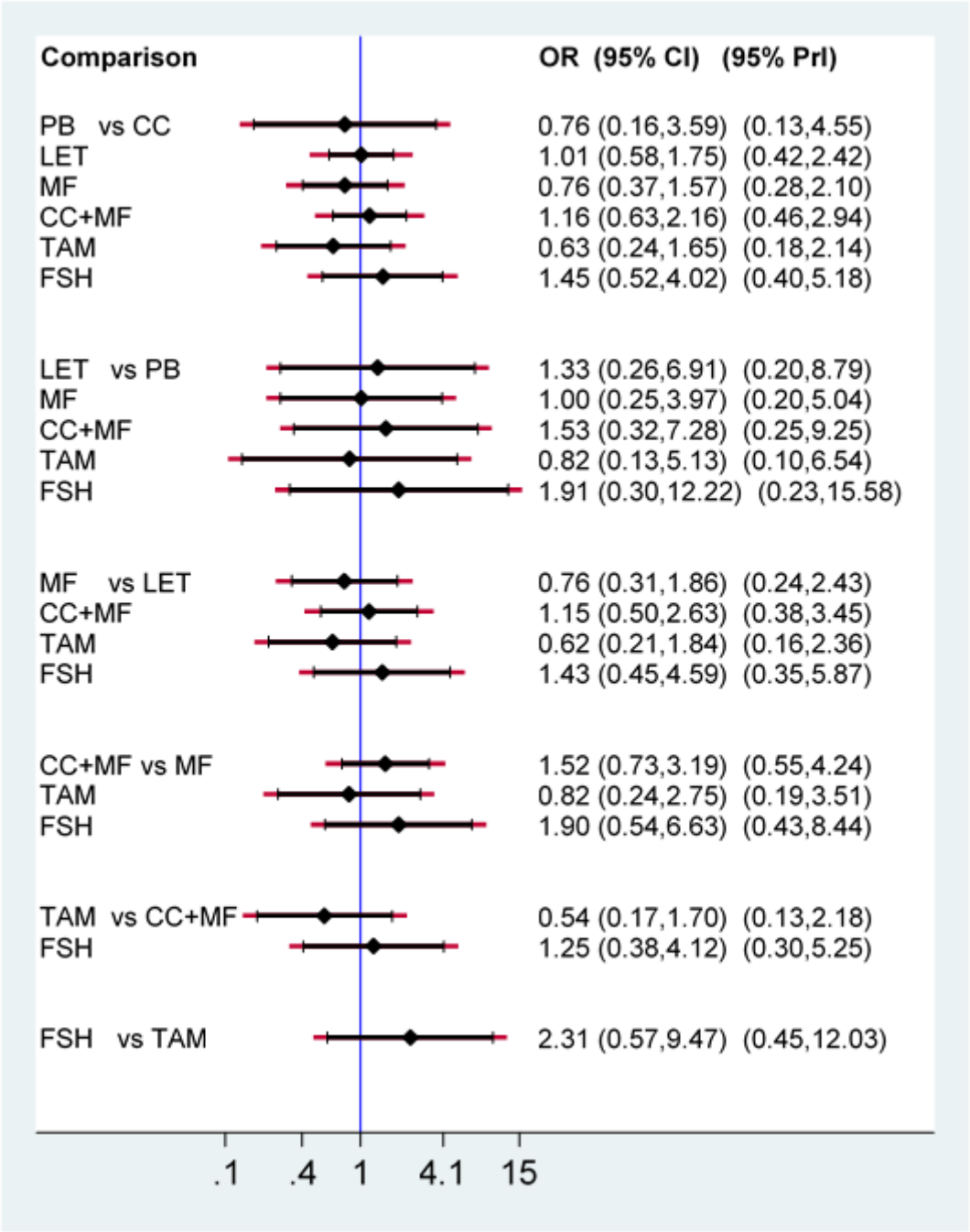
22a. Risk of bias in randomisation



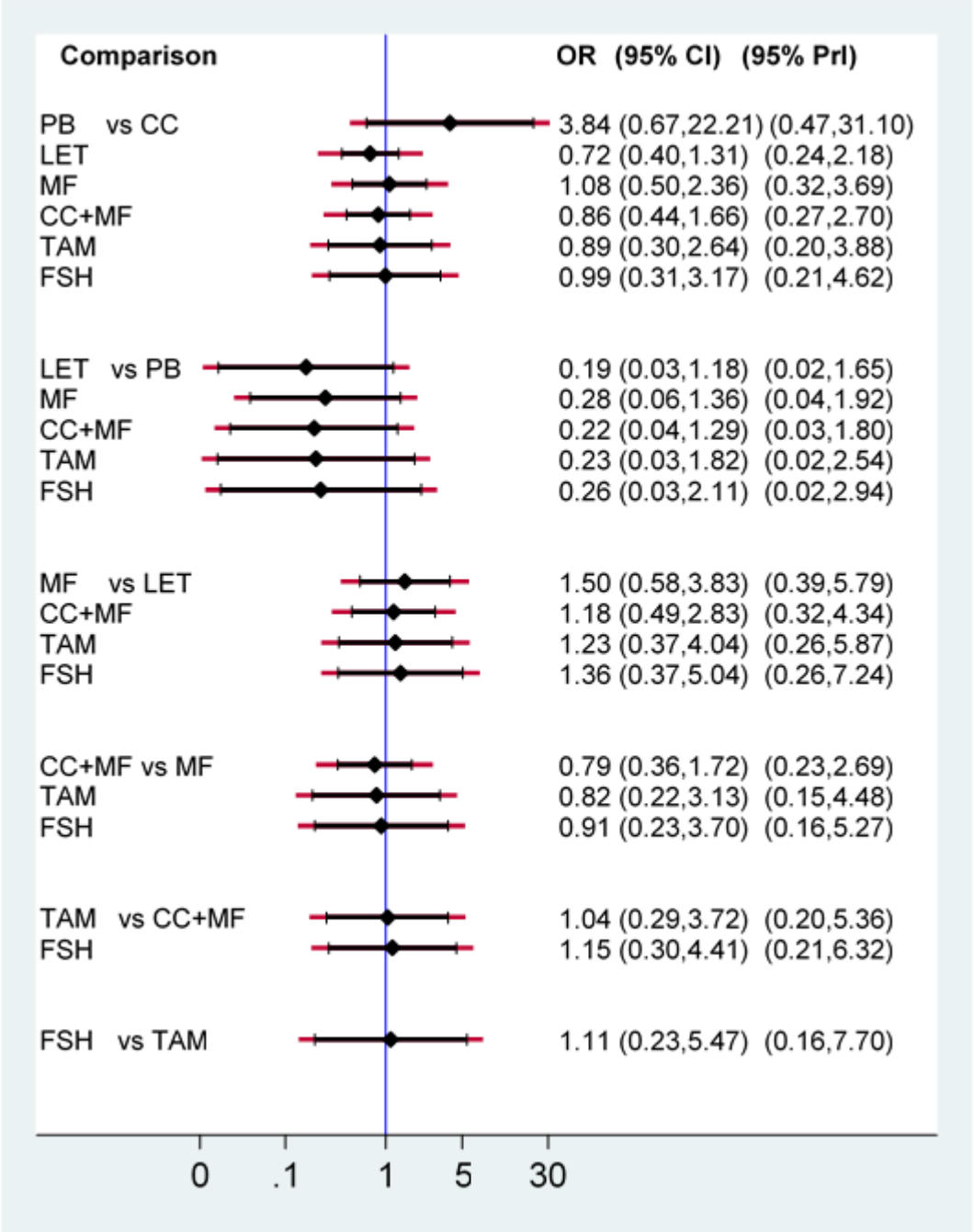
22b. Risk of bias in allocation concealment



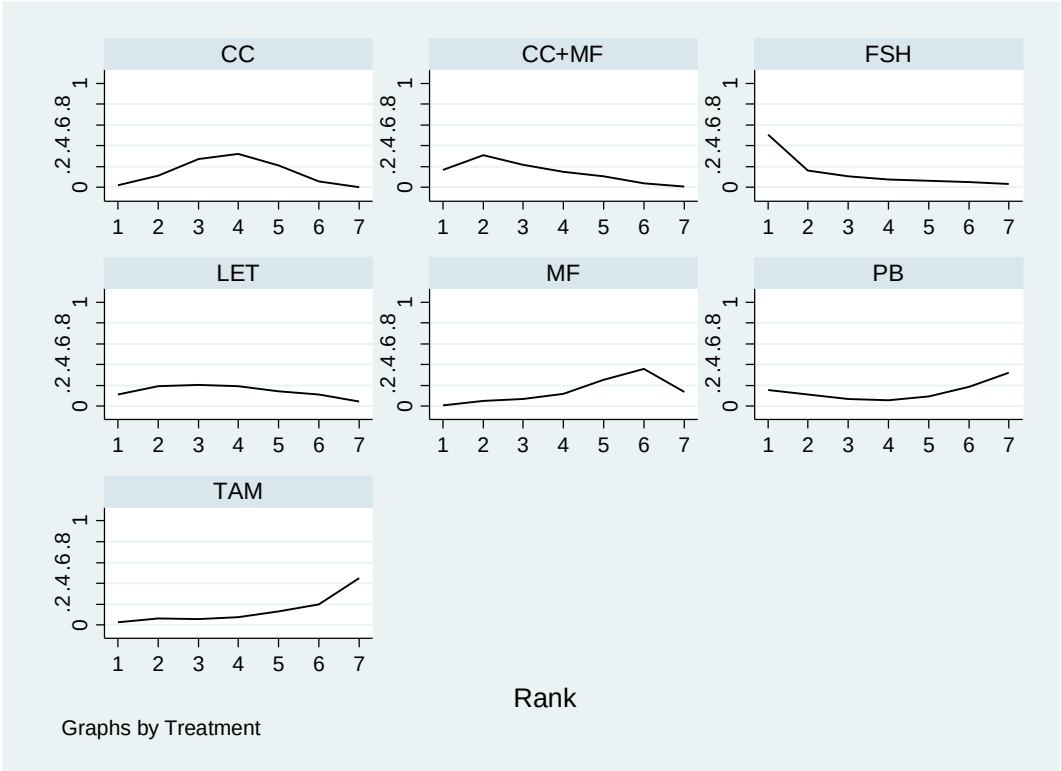
Appendix 23 Network meta-analysis results for miscarriage per woman randomised



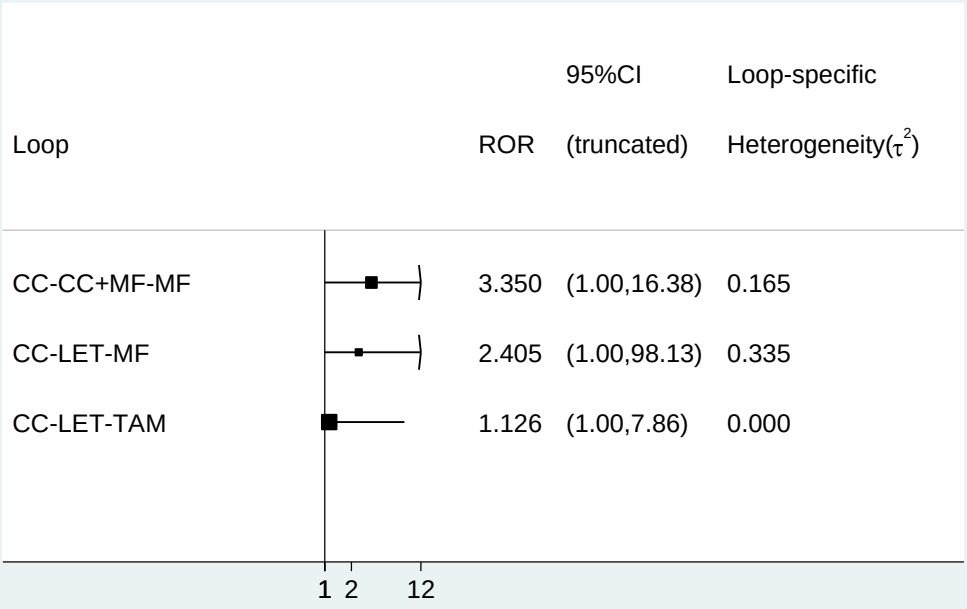
Appendix 24 Network meta-analysis results for miscarriage per pregnancy



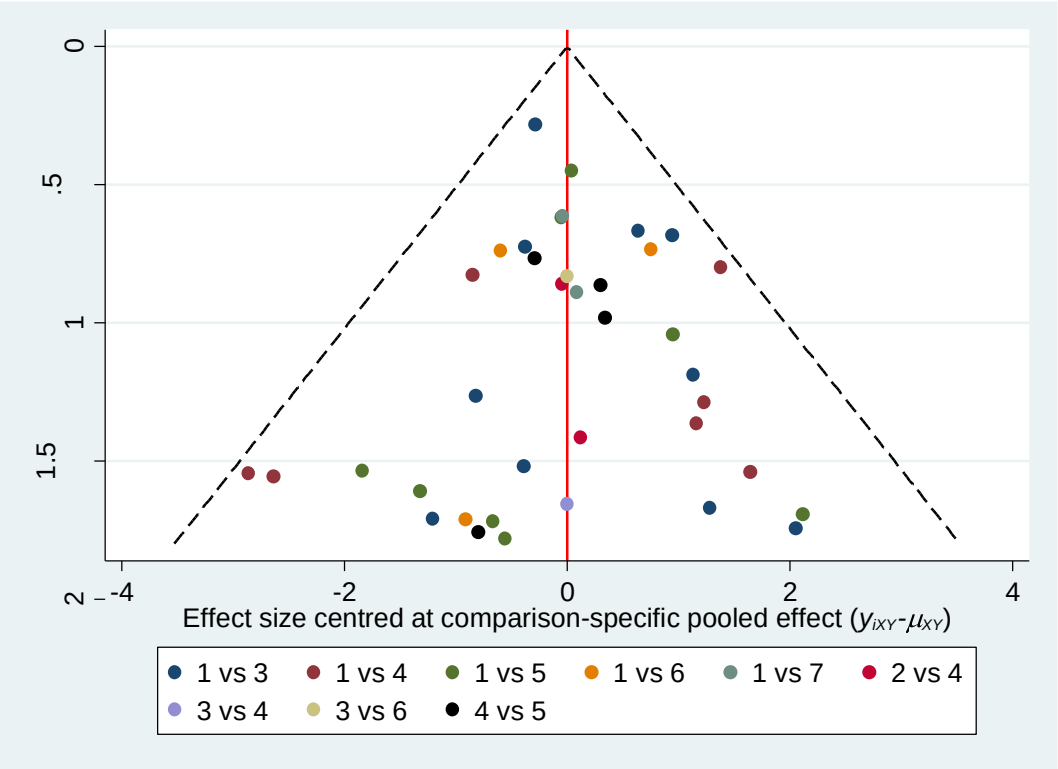
Appendix 25 Ranking of treatments for miscarriage per pregnancy



Appendix 26 Inconsistency plot for miscarriage per pregnancy

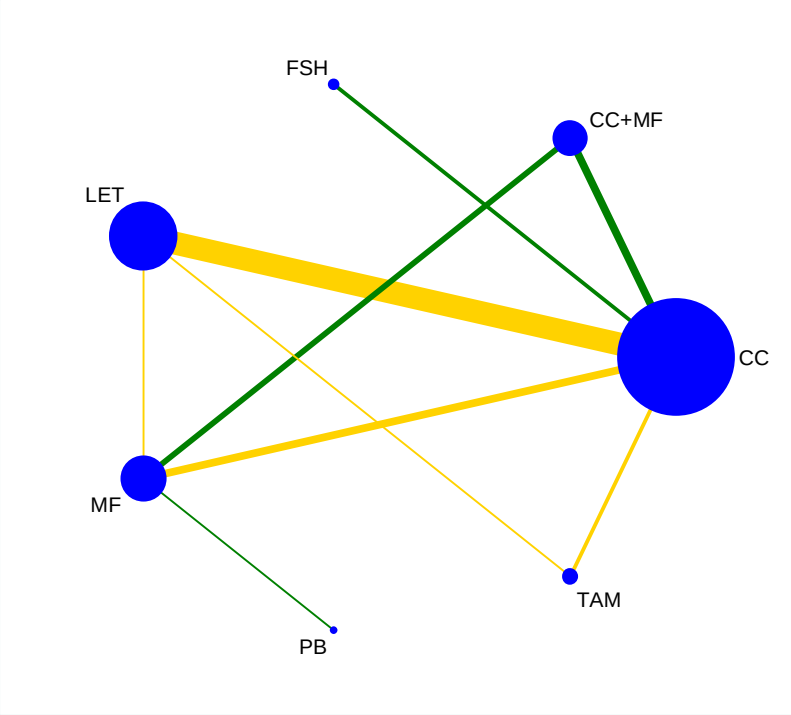


Appendix 27 Comparison-adjusted funnel plot for miscarriage per pregnancy

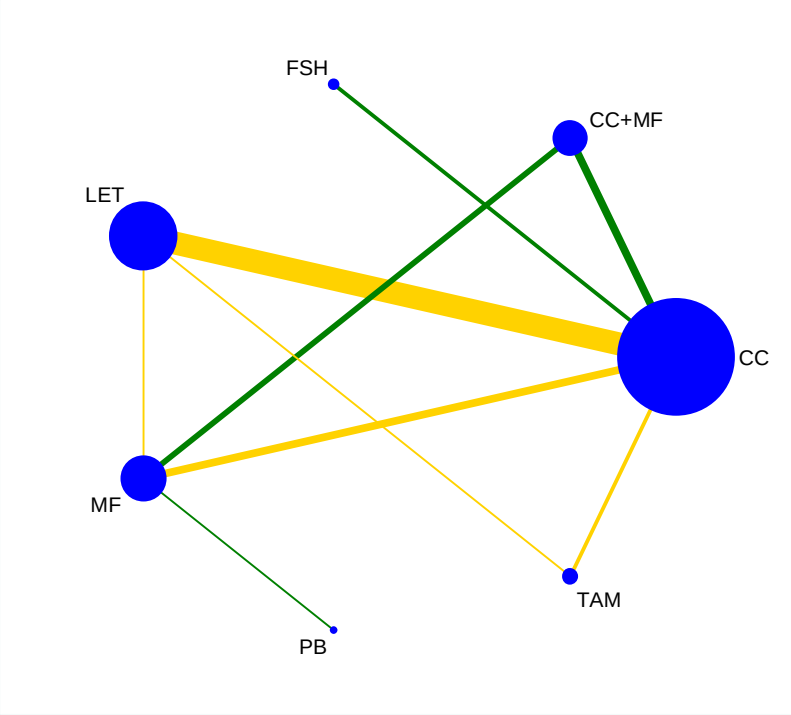


(1-clomiphene; 2-placebo/no treatment; 3-letrozole; 4-metformin; 5-clomiphene plus metformin; 6-tamoxifen; 7-FSH)

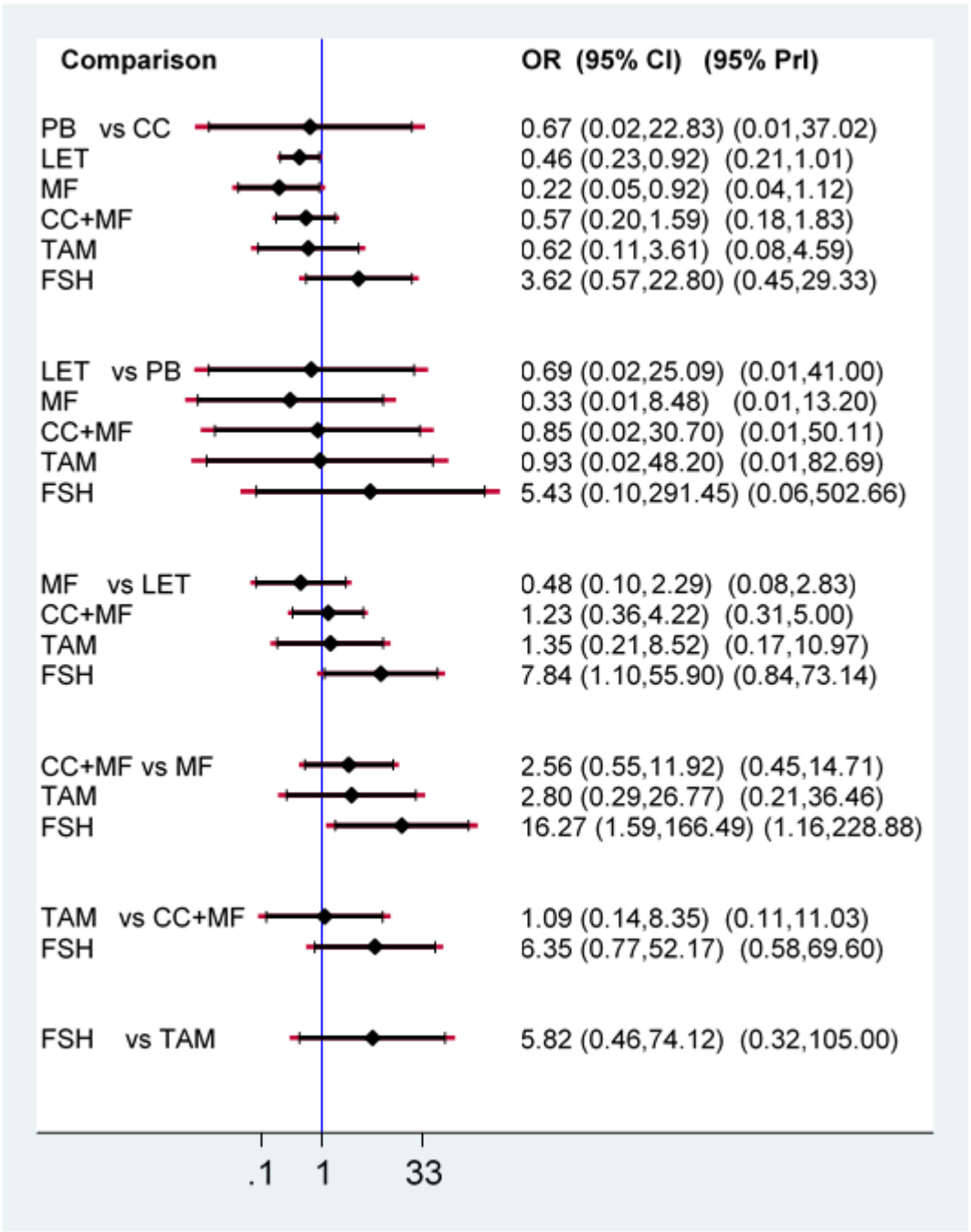
Appendix 28 Network plot for multiple pregnancy incorporating risk of bias assessment
28a. Risk of bias in randomisation



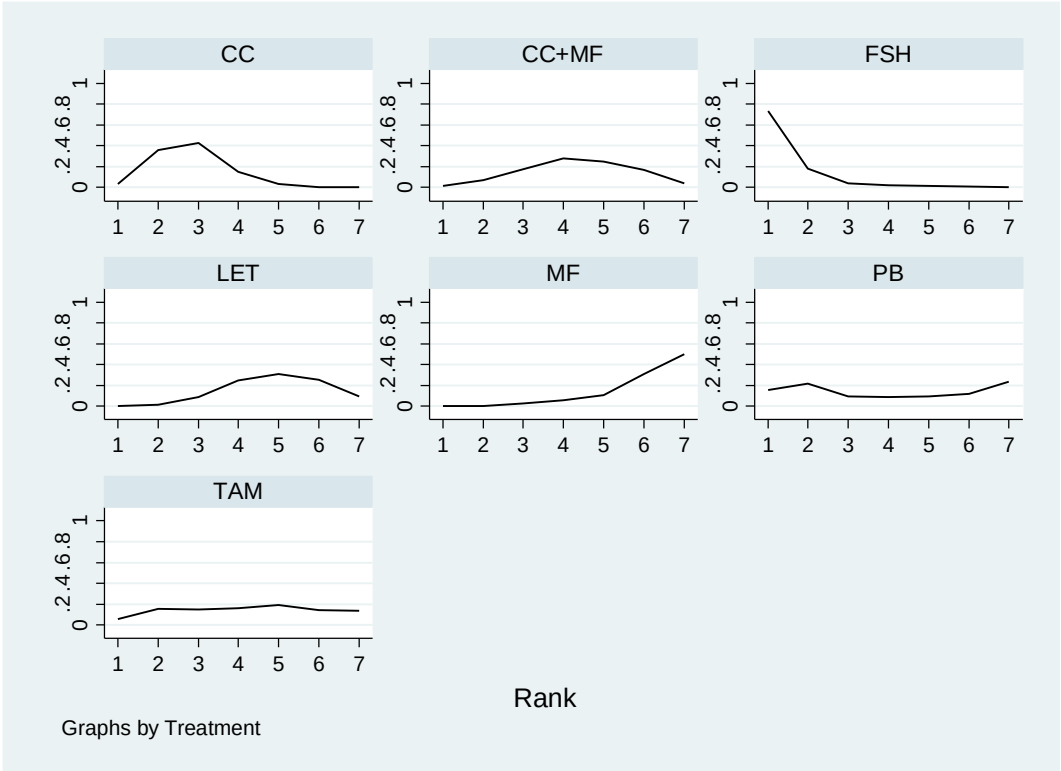
28b. Risk of bias in allocation concealment



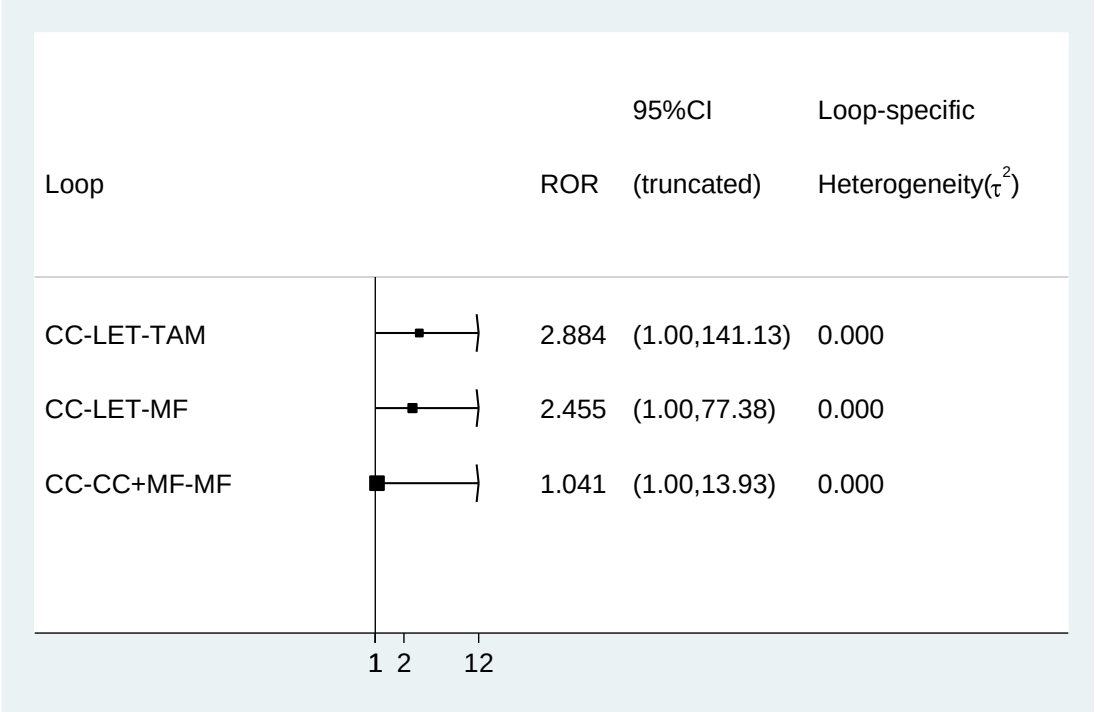
Appendix 29 Network meta-analysis results for multiple pregnancy



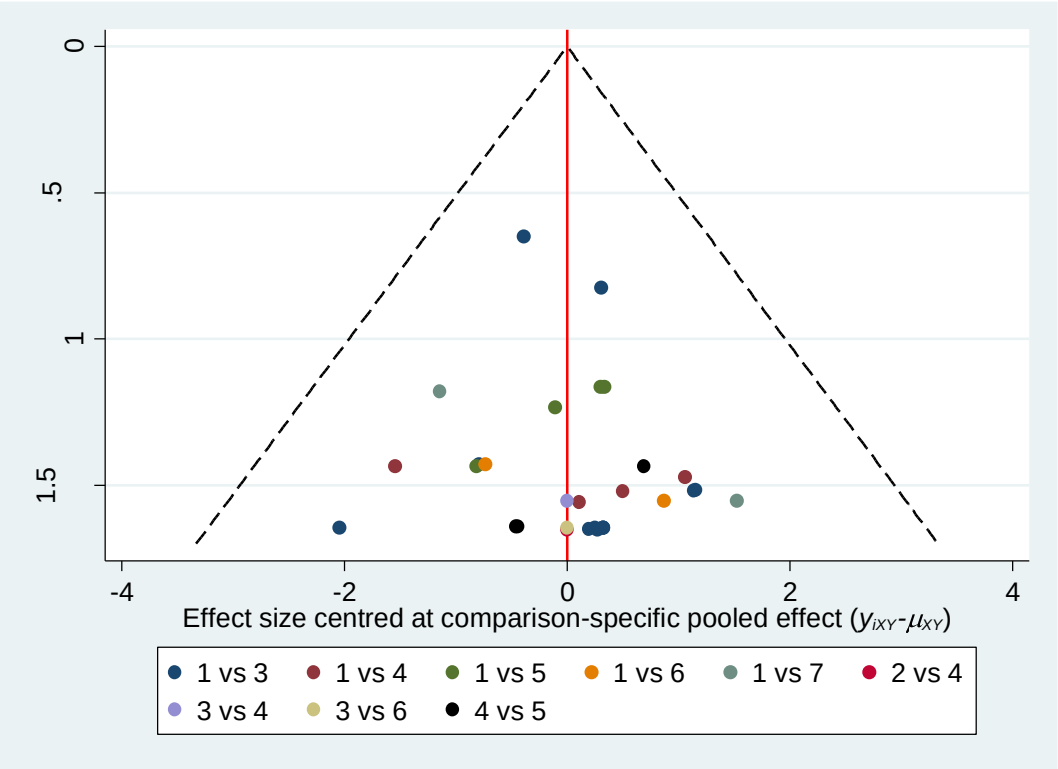
Appendix 30 Ranking of treatments for multiple pregnancy



Appendix 31 Inconsistency plot for multiple pregnancy

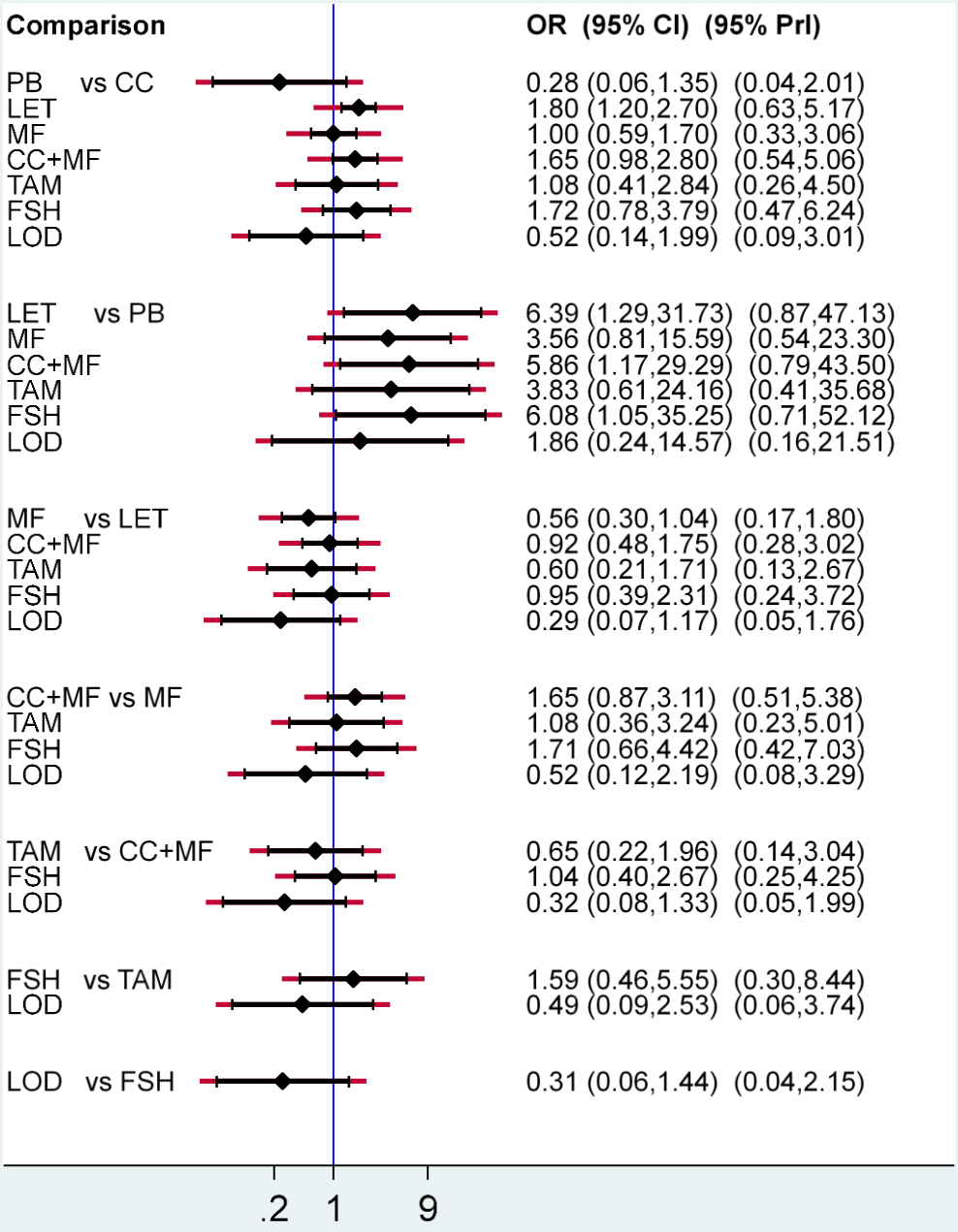


Appendix 32 Comparison-adjusted funnel plot for multiple pregnancy

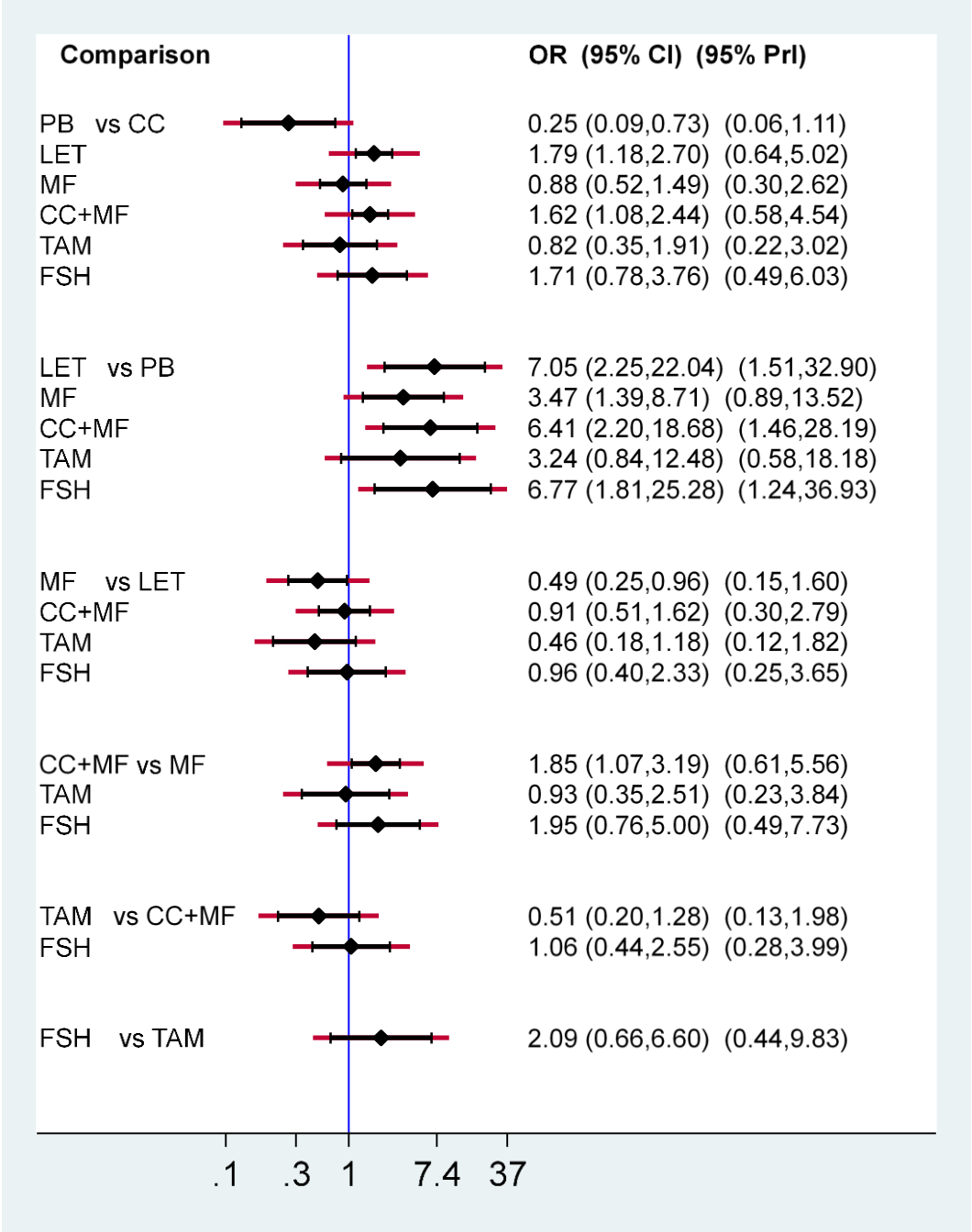


(1-clomiphene; 2-placebo/no treatment; 3-letrozole; 4-metformin; 5-clomiphene plus metformin; 6-tamoxifen; 7-FSH)

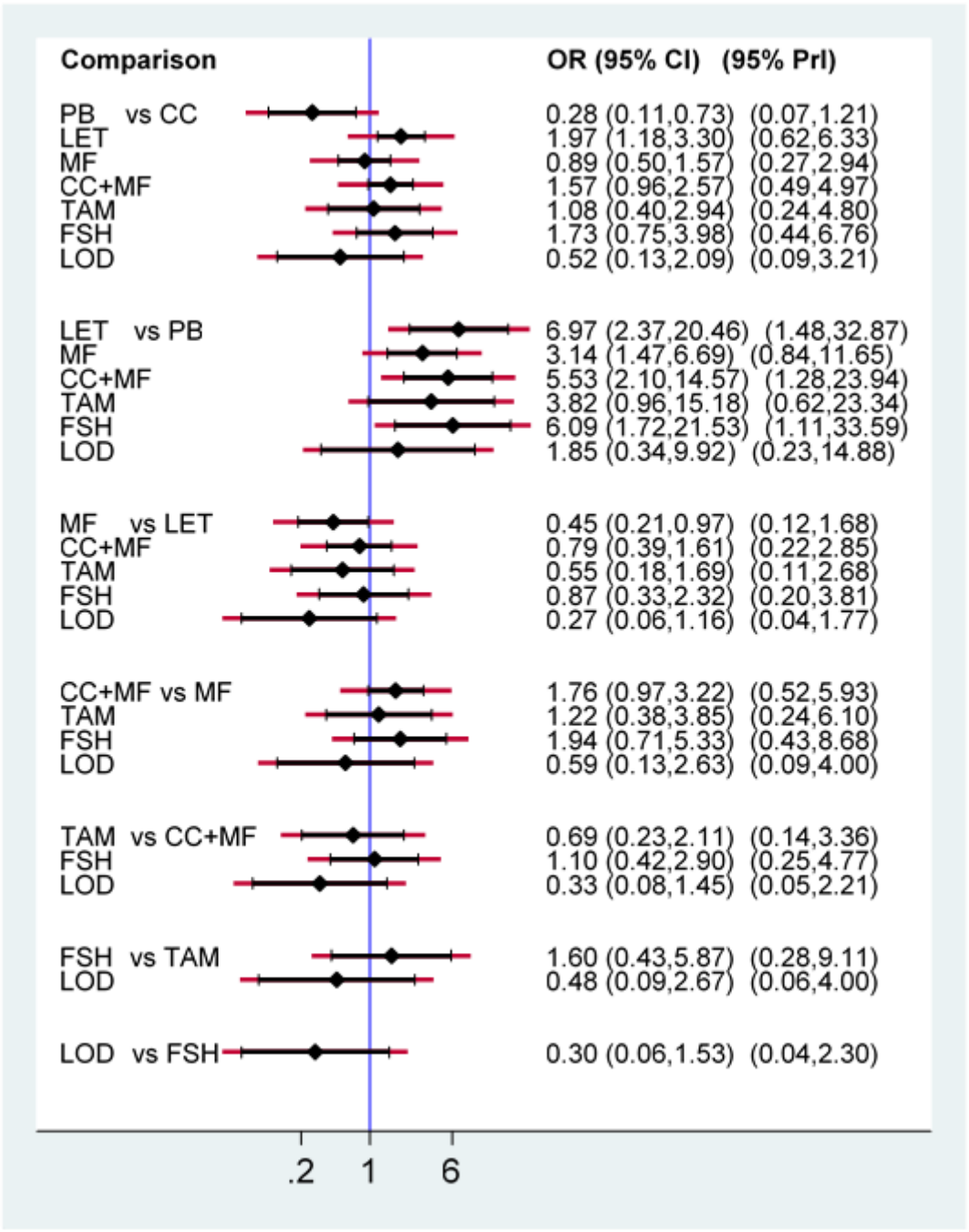
Appendix 33 Sensitivity analysis - RCTs with treatment naïve women



Appendix 34 Sensitivity analysis - RCTs reporting clinical pregnancy



Appendix 35 Sensitivity analysis - RCTs with low risk of randomisation & allocation bias



Appendix 36 Additional discussion

Side effects of the combination of clomiphene and metformin

We have summarised the side effects of the combination of clomiphene and metformin versus clomiphene alone in a supplementary table (Appendix 38). Of the 19 studies comparing these two interventions, 11 studies reported data on side effects or discontinuation due to side effects. Three studies[1-3] including 714 women reported the number of participants who discontinued treatment due to side effects. In a pairwise meta-analysis for this outcome, we found that more women in the combination group discontinued the treatment due to side effects than women in clomiphene group (OR 2.34, 95% CI 1.04 to 5.30, Appendix 37). As the reporting strategies were diverse in different studies, we were not able to perform meta-analyses on overall side effects or any specific types of side effects. As shown in Appendix 38, gastrointestinal side effects were more frequent in combined clomiphene-metformin group than clomiphene group.

Quality of evidence and interpretation of data

The overall quality of included studies was moderate in relation to the seven specific domains of the risk of bias assessment. Randomisation and allocation are fundamental requirements for a high quality RCT and therefore we integrated these domains in the network plot (Appendix 8, 12, 17, 22, 28). Although we excluded quasi-randomised studies in the current systematic review, half of the included RCTs did not report details of randomisation, and further clarity on this eluded us even after attempts to contact the authors. Specific information about allocation concealment was also unavailable in many of the trials. In multicentre RCTs with large sample sizes[1 2 4 5], the dropout rates in different interventions varied from 14% to 35%. Many studies with small sample sizes have relatively low or zero dropout rates. Additionally, these studies often claim to have undertaken an intention-to-treat analysis, but it is possible that the authors may have excluded dropouts in their analysis. It is difficult to distinguish those lost to follow up due to adverse events and those for other reasons. CONSORT[6] strongly encourages to report a flow diagram of patient follow up, including reasons for dropouts, however, many included studies failed to do so.

In pairwise meta-analyses, the heterogeneity in comparisons of combined clomiphene-metformin versus clomiphene and letrozole versus clomiphene in all outcomes was low. Therefore, the results of these comparisons in network meta-analysis were robust. By contrast, there was significant heterogeneity in comparisons of clomiphene and metformin. Thus, the results of these comparisons should be interpreted with cautions.

In our network meta-analysis, predictive intervals were used to estimate the effect of a future study. When considering predictive interval in our network meta-analysis, clomiphene, letrozole, metformin, combined clomiphene- metformin, and FSH remained superior to placebo. These results indicate that in future studies, these active treatments would remain effective in comparison with placebo/no treatment. Of note, there were significant differences between FSH and metformin/letrozole in terms of multiple pregnancy. However, the wide confidence intervals suggest significant imprecision in the effect size.

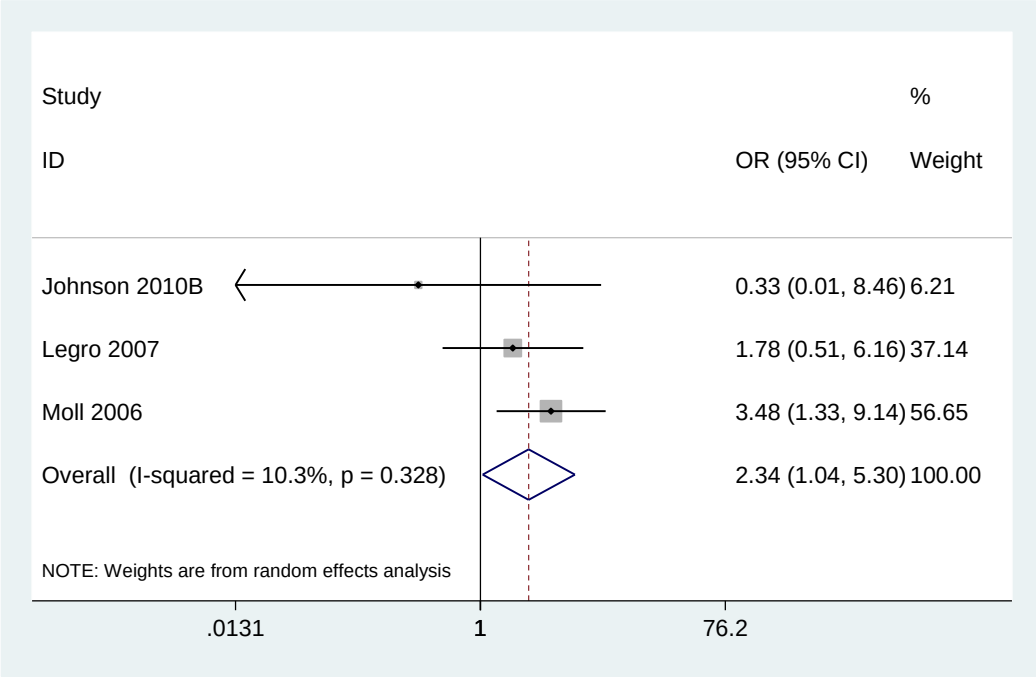
According to the rankings, combined clomiphene-metformin, letrozole, and FSH were the best interventions in terms of pregnancy, live birth and ovulation, while metformin and letrozole were the best interventions in terms of reducing multiple pregnancy rate.

References

1. Legro RS, Barnhart HX, Schlaff WD, et al. Clomiphene, metformin, or both for infertility in the polycystic ovary syndrome. New England journal of medicine 2007;**356**(6):551-66.
2. Moll E, Bossuyt PM, Korevaar JC, et al. Effect of clomifene citrate plus metformin and clomifene citrate plus placebo on induction of ovulation in women with newly diagnosed polycystic ovary syndrome: randomised double blind clinical trial. BMJ (Clinical research ed)

2006;**332**(7556):1485.
3. Johnson NP, Stewart AW, Falkiner J, et al. PCOSMIC: a multi-centre randomized trial in women with PolyCystic Ovary Syndrome evaluating Metformin for Infertility with Clomiphene. Human reproduction 2010;**25**(7):1675-83.
4. Legro RS, Brzyski RG, Diamond MP, et al. Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. The New England journal of medicine 2014;**371**(2):119-29.
5. Homburg R, Hendriks ML, Konig TE, et al. Clomifene citrate or low-dose FSH for the first-line treatment of infertile women with anovulation associated with polycystic ovary syndrome: a prospective randomized multinational study. Human reproduction 2012;**27**(2):468-73.
6. Moher D, Schulz KF, Altman D, et al. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. Jama 2001;**285**(15):1987-91.

Appendix 37 Meta-analysis of combined clomiphene-metformin vs clomiphene for discontinuation due to side effects.



Boxes and horizontal lines represent ORs and 95% CIs of individual studies. The diamond represents the overall OR and 95% CI (Random-effect model). OR >1 means more women discontinue treatment due to side effect in combined clomiphene-metformin group than clomiphene group.

Appendix 38 Side effects of combined clomiphene-metformin versus clomiphene alone.

Study ID	CC group				CC+Metformin group			
	Women with side effects	Details of side effects	discontinuation due to side effects	Sample size	Women with side effect	Details of side effect	discontinuation due to side effects	Sample size
Abuelghar 2013	6	Flushing: 4; gastrointestinal tract discomfort: 2	NA	32	11	Flushing: 2; gastrointestinal tract discomfort: 5; both: 1; diarrhoea: 3	NA	34
Ayaz 2013	NA	NA	0	21	NA	60% had complained of loss of appetite, 18% had nausea & vomiting	0	21
Basirat 2012	NA	No metformin related side effects.	NA	167	NA	No metformin related side effects.	NA	167
Dasari 2009	NA	NA	0	24	NA	NA[a]	0	16
Johnson 2010B	NA	Gastrointestinal symptoms: 5	1	36	NA	Gastrointestinal symptoms: 11; vasomotor: 1	0	35
Legro 2007	NA	Diarrhoea: 48; dyspepsia: 9; flatulence: 38; nausea: 82; stomach discomfort: 8; vomiting: 28; decreased appetite: 17[b]	4	209	NA	Diarrhoea: 126; dyspepsia: 14; flatulence: 30; nausea: 138; stomach discomfort: 16; vomiting: 72; decreased appetite: 33 [b]	7	209
Maged 2015	1	Nausea: 1	NA	40	1	Drowsiness:1	NA	40
Moll 2006	NA	NA	6	114	NA	NA	18	111
Raja 2005	NA	NA	NA	50	6	nausea and diarrhoea: 6	NA	50
Sahin 2004	NA	NA	0	10	NA	NA	0	11
Zain 2009	NA	NA	0	41	NA	NA[c]	0	41

NA: not available.

a. The data of the 16 women in CC+ metformin group were not reported. But the authors reported that of the 25 participants who received metformin along with CC, 80% complained of loss of appetite and 24% had nausea and vomiting. The 25 participants was composed of 16 women in CC + metformin group and 9 women who did not conceive with six cycles of CC alone (given CC + metformin for an additional six cycles) at their request for further treatment.

b. Main gastrointestinal side effects were summarised in this table. This study also reported data on other specific side effects but not the data on overall side effects.

c. The data of CC+metformin group was not reported. Three patients with metformin complained of nausea, dizziness, and headache.

Appendix 39 Congenital malformations in newborns conceived through letrozole vs control.

Study ID	Country	Study design	Congenital malformation	
			Control	**Letrozole**
Dehbashi 2009[1]	Iran	RCT	CC: 16.6% (1/6) [a]	0% (0/10)
Ray 2012[2]	India	RCT	CC: 0% (0/13)	0% (0/20)
Roy 2012[3]	India	RCT	CC: 0% (0/21)	0% (0/39)
Legro 2014[4]	USA	RCT	CC: 1.5% (1/66) [b]	3.9% (4/102) [c]
Diamond 2015[5]	USA	RCT	CC: 4.3% (3/70) [d]	3.6% (2/56) [e]
Tulandi 2006[6]	Canada	observational	CC/CC+FSH: 4.8(19/397) [f]	Letrozole/Letrozole+FSH: 2.4% (14/514) [g]
Forman 2007[7]	Canada	observational	2.6% (7/271) [h]	0% (0/94)
Sharma 2014[8]	India	observational	CC:4.0% (10/251) [I]; Natural conception: 2.9% (5/171)[k]	2.5% (5/201) [j]
Wu 2016[9]	China	RCT	Berberine: 0% (0/48)	Letrozole alone: 1.2%(1/84)[l]; Letrozole+Berberine: 1.2%(1/81)[m]
Tatsumi 2016[10]	Japan	observational	Natural cycle IVF/ICSI: 1.9% (44/2287)[n]	Letrozole + IVF/ICSI: 2.2%(15/694)[o]

Details of congenital malformations in these studies:

a. Meningomyelocele.

b. Atrial septal defect (ASD), ventricular septal defect (VSD), and pulmonary stenosis.

c. 1) Cerebral palsy with arrested hydrocephalus with polycythemia and neutropenia; 2) imperforate anus with perineal fistula and spina bifida with a tethered spinal cord; 3) right hemimegalencephaly, and dysgenesis of the left frontal and temporal lobes but no hydrocephalus; 4) large cardiac VSD requiring surgical repair.

d. 1) Aortic arch hypoplasia; 2) Congenital hypothyroidism; 3) Renal duplicated right collecting system and ureterocele.

e. 1) Hypospadias; 2)Right facial hemangioma; Biventricular hypertrophy; Bifid uvula; Small cataracts bilaterally; Widening of the corneal horizontal diameter.

f. Major malformations (12 cases): 1) VSD (4 cases); 2) Transposition of great vessels; 3) Atresia of pulmonary valve and right ventricle; 4) Pulmonary valve atresia; 5) Pyelectasis; 6) Omphalocele; 7) Cleft palate; 8) Spinal muscular atrophy; 9) Down's syndrome.
Minor malformations (7 cases): 1) Preauricular skin tag (2 cases); 2) Horseshoe kidney; 3) Polydactyly (3 cases); 4) Unspecific hypotonia.

g. Major malformations (6 cases): 1) VSD; 2) Esophageal atresia; 3) Cleft palate; 4) Trisomy 18; 5) Down's syndrome; 6) Potter's syndrome.
Minor malformations (8 cases): 1) Preauricular skin tag; 2) Congenital ptosis; 3) Plagiocephaly; 4) Hydrocele; 5) Hypospadia; 6) Polydactyly; 7) Syndactyly (2nd and 3rd toes); 8) Umbilical and inguinal hernias.

h. 7 cases with major malformations but details not reported.

i. 1) Patent ductus arteriosus (2 cases) and; 2) total anomalous venous connection; 3) Hypospadias (3 cases); 4) bilateral congenital talipus equino varus; 5) duplication of urethra; 6) cleft lip & palate; 7)inguinal hernia; 8)neural tube defect; 9) Down's syndrome (2 cases). Three babies with congenital heart disease were excluded from the analysis by the authors as they were born to diabetic mothers.

j. 1) Combined ventricular and ASD; 2) paraumbilical hernia; 3) congenital deafness; 4) congenital talipus equino varus; 5) albinism.

k. 1) VSD; 2) Congenital talipus equino varus; 3)cleft lip; 4) imperforate anus; 5)polydactyly.

l. Hydrocephalus.

m. Major VSD and pulmonary stenosis.

n. Major anomalies (34 cases): including chromosomal abnormalities (11 cases), cardiovascular abnormalities (13 cases) and musculoskeletal abnormalities (1 case).
o. Major anomalies (13 cases): 1) ASD,VSD; 2)ASD,VSD, Down's syndrome; 3) Cleft lip without cleft palate; 4) Congenital hydronephrosis; 5) Diaphragmatic hernia; 6) Duodenal atresia; 7) Endocardial cushion defect, down syndrome; 8) Hypospadias; 9) Trisomy 18; 10) VSD (2 cases); 11)VSD, down syndrome; 12) Anencephalus.

References:

1. Dehbashi S, Dehbashi S, Kazerooni T, et al. Comparison of the effects of letrozole and clomiphene citrate on ovulation and pregnancy rate in patients with polycystic ovary syndrome. Iranian Journal of Medical Sciences 2009;**34**(1):23-28.
2. Ray PB, Ray A, Chakraborti PS. Comparison of efficacy of letrozole and clomiphene citrate in ovulation induction in Indian women with polycystic ovarian syndrome. Archives of Gynecology & Obstetrics 2012;**285**(3):873-7.
3. Roy K, Baruah J, Singla S, et al. A prospective randomized trial comparing the efficacy of Letrozole and Clomiphene citrate in induction of ovulation in polycystic ovarian syndrome. Journal of human reproductive sciences 2012;**5**(1):20-25.
4. Legro RS, Brzyski RG, Diamond MP, et al. Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. The New England journal of medicine 2014;**371**(2):119-29.
5. Diamond MP, Legro RS, Coutifaris C, et al. Letrozole, Gonadotropin, or Clomiphene for Unexplained Infertility. The New England journal of medicine 2015;**373**(13):1230-40.
6. Tulandi T, Martin J, Al-Fadhli R, et al. Congenital malformations among 911 newborns conceived after infertility treatment with letrozole or clomiphene citrate. Fertility and sterility 2006;**85**(6):1761-5.
7. Forman R, Gill S, Moretti M, et al. Fetal safety of letrozole and clomiphene citrate for ovulation induction. J Obstet Gynaecol Can 2007;**29**(8):668-71.
8. Sharma S, Ghosh S, Singh S, et al. Congenital malformations among babies born following letrozole or clomiphene for infertility treatment. PloS one 2014;**9**(10):e108219.
9. Wu XK, Wang YY, Liu JP, et al. Randomized controlled trial of letrozole, berberine, or a combination for infertility in the polycystic ovary syndrome. Fertility and sterility 2016;**106**(3):757-65 e1.
10. Tatsumi T, Jwa SC, Kuwahara A, et al. No increased risk of major congenital anomalies or adverse pregnancy or neonatal outcomes following letrozole use in assisted reproductive technology. Human reproduction 2016.

Supplemental materials for Chapter 4

Supplemental Table 1 Search strategy

Database	Search strategy
Medline	#1. polycystic ovary syndrome.sh #2. polycystic ovar*.ti,ab,kw #3. ovary polycystic disease.ti,ab,kw #4. anovulation.sh #5. anovulat*.ti,ab,kw #6. oligoovulat*.ti,ab,kw #7. (oligo ADJ2 ovulat*).ti,ab,kw #8. PCOS. ti,ab,kw #9. PCOD.ti,ab,kw #10. stein leventhal syndrome.ti,ab,kw #11. sclerocystic ovar*.ti,ab,kw #12. ovarian degeneration.ti,ab,kw #13. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 #14. Clomiphene.sh #15. clomiphene.ti,ab,kw #16. clomifene.ti,ab,kw #17. clomid.ti,ab,kw #18. Serophene.ti,ab,kw #19. Gravosan.ti,ab,kw #20. #14 OR #15 OR #16 OR #17 or #18 OR #19 #21. ((randomized controlled trial or controlled clinical trial).pt. or (randomi* or placebo* or randomly).mp. or drug therapy.fs.) #22. #13 AND #20 AND #21
Embase	#1. 'ovary polycystic disease'/syn #2. 'ovary polycystic disease':ti,ab,de #3. 'polycystic ovar*':ti,ab,de #4. 'anovulation'/syn #5. anovulat*:ti,ab,de #6. oligoovulat*:de,ab,ti #7. ((oligo NEAR/2 ovulat*):de,ab,ti) #8. PCOS:ti,ab,de #9. PCOD:ti,ab,de #10. 'stein leventhal syndrome':ti,ab,de #11. 'sclerocystic ovar*':ti,ab,de #12. 'ovarian degeneration':ti,ab,de #13. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 #14. 'clomifene'/syn #15. clomifene:ti,ab,de #16. clomiphene:ti,ab,de #17. clomid:ti,ab,de #18. Serophene:ti,ab,de #19. Gravosan:ti,ab,de #20. #14 OR #15 OR #16 OR #17 OR #18 OR #19 #21. (('randomized controlled trial'/syn OR randomi* NEXT/10 controlled OR randomization/syn OR 'random allocation' OR 'double blind procedure'/syn OR (single OR double OR tripl* OR trebl*) NEXT/1 (blind* OR mask*) OR 'clinical trial'/syn OR 'clinical trials' OR 'multicenter study'/syn OR (multicentre OR multicenter) NEXT/1 stud* OR randomly:ti,ab OR trial:ti,ab OR groups:ti,ab) NOT (animal/syn NOT human/syn)) #22. #13 AND #20 AND #21

CENTRAL	#1. MeSH descriptor: [Polycystic Ovary Syndrome] explode all trees #2. polycystic ova*:ti,ab,kw (Word variations have been searched) #3. ovary polycystic disease:ti,ab,kw (Word variations have been searched) #4. PCOS:ti,ab,kw (Word variations have been searched) #5. PCOD:ti,ab,kw (Word variations have been searched) #6. MeSH descriptor: [Anovulation] explode all trees #7. anovulat*:ti,ab,kw (Word variations have been searched) #8. oligoovulat*:ti,ab,kw (Word variations have been searched) #9. oligo near/2 ovulat*:ti,ab,kw (Word variations have been searched) #10. 'Stein Leventhal':ti,ab,kw (Word variations have been searched) #11. ovarian degeneration:ti,ab,kw (Word variations have been searched) #12. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 #13. MeSH descriptor: [Clomiphene] explode all trees #14. clomiphene:ti,ab,kw (Word variations have been searched) #15. clomifene:ti,ab,kw (Word variations have been searched) #16. clomid:ti,ab,kw (Word variations have been searched) #17. serophene:ti,ab,kw (Word variations have been searched) #18. #13 OR #14 OR #15 OR #16 OR #17 #19. #12 AND #18 #20. Add filter: trials

Supplemental Table 2 Excluded studies and reasons for exclusion

Reasons for exclusion	References
Duplicate population: n=6;	1. Al Boghdady L, Ghanem ME, Hassan M, Helal AS. Clomiphene citrate (CC) co-treatment with low dose urinary FSH versus. Urinary FSH protocol for CC resistant PCOS: Randomized controlled trial. Human Reproduction 2012;27. 2. Foroozanfard FMMMG. Comparing letrozole and clomiphene in combined regimens with gonadotropins in pregnancy rate in patients with clomiphen resistant polycystic ovarian syndrome. In. Vol. 8, 2010:15 Abstract no: O-32. 3. Javdani M, Fallahzadeh H, Davar R, Sheibani H. Metformin-letrozole in comparison with metformin-clomiphene citrate in clomiphene-resistance pco patients undergoing IUI. Human Reproduction 2012;27. 4. Palomba S, Orio Jr F, Falbo A, Russo T, Caterina G, Manguso F et al. Metformin administration and laparoscopic ovarian drilling improve ovarian response to clomiphene citrate (CC) in oligo-anovulatory CC-resistant women with polycystic ovary syndrome. Clinical Endocrinology 2005;63:631-5. 5. Weiss N, Nahuis M, Bordewijk E, Oosterhuis J, Lambalk C, Koks C et al. Anovulatory women not conceiving after six ovulatory cycles with clomiphene citrate-should we switch to gonadotrophins and/or add IUI? A 2 by2 factorial RCT. Human Reproduction 2017;32:i6. 6. Yarali H, Yyldyz B, Demirol A, Zeyneloglu H, Yigit N, Bukulmez O. Co-administration of metformin during recombinant follicle stimulating hormone (recombinant FSH) treatment using the low-dose step protocol in patients with clomiphene citrate resistant polycystic ovary syndrome (PCOS): a prospective randomized trial. In. Vol. 76, 2001:S36.
Not an RCT: n = 7;	1. Chen ZJ, Shi YH, Li Y, Gao Q, Sheng Y, Ma ZX. [Clinical analysis of assistant treatment proposals for infertile women with polycystic ovary syndrome]. Zhonghua fu chan ke za zhi 2008;43:571-5. 2. Gadir AA, Mowafi RS, Alnaser HMI, Alrashid AH, Alonezi OM, Shaw RW. Ovarian electrocautery versus human menopausal gonadotrophins and pure follicle stimulating hormone therapy in the treatment of patients with polycystic ovarian disease. Clinical Endocrinology 1990;33:585-92. 3. Kazerooni T, Ghaffarpasand F, Kazerooni Y, Kazerooni M, Setoodeh S. Short-term metformin treatment for clomiphene citrate-resistant women with polycystic ovary syndrome. International Journal of Gynaecology & Obstetrics 2009;107:50-3. 4. Kocak M, Caliskan E, Simsir C, Haberal A. Metformin therapy improves ovulatory rates, cervical scores, and pregnancy rates in clomiphene citrate-resistant women with polycystic ovary syndrome. Fertility & Sterility 2002;77:101-6. 5. Salaheldin AbdelHamid AM, Rateb AM, Ismail Madkour WA. Is clomiphene citrate stair-step protocol a good alternative to gonadotrophins in clomiphene-resistant PCO patients? Prospective study. Journal of Obstetrics & Gynaecology Research 2016;42:547-53. 6. Sharma S, Rani G, Bose G, Saha I, Bathwal S, Chakravarty B. Tamoxifen is better than low-dose clomiphene or gonadotropins in women with thin endometrium (<7 mm) after clomiphene in intrauterine insemination cycles: A prospective study. Journal of Human Reproductive Sciences 2018;11:34-9. 7. Xi W, Liu S, Mao H, Yang Y, Xue X, Lu X. Use of letrozole and clomiphene citrate combined with gonadotropins in clomiphene-resistant infertile women with polycystic ovary syndrome: a prospective study. Drug design, development & therapy 2015;9:6001-8.
Irrelevant population: n=4;	1. Behnoud N, Farzaneh F, Ershadi S. The effect of clomiphene citrate versus letrozole on pregnancy rate in women with polycystic ovary syndrome: A randomized clinical trial. Crescent Journal of Medical and Biological Sciences 2019;6:335-40.

	2. Fatima A, Khan SA, Saifuddin Z, Aslam R. Comparison of efficacy of clomiphene citrate alone and with metformin for treatment of infertility in polycystic ovarian syndrome. Rawal Medical Journal 2018;43:285-8. 3. Mejia RB, Summers KM, Kresowik JD, Van Voorhis BJ. A randomized controlled trial of combination letrozole and clomiphene citrate or letrozole alone for ovulation induction in women with polycystic ovary syndrome. Fertility and Sterility 2019;111:571-8.e1. 4. Wang L, Wen X, Lv S, Zhao J, Yang T, Yang X. Comparison of endometrial receptivity of clomiphene citrate versus letrozole in women with polycystic ovary syndrome: a randomized controlled study. Gynecological Endocrinology 2019.
Irrelevant interventions: n=4;	1. Badawy AM, Allam A, Abulatta M. Extending clomiphene treatment in clomiphene-resistant women with PCOS: A randomized controlled trial. Reproductive Biomedicine Online 2008;16:825-9. 2. Kaya H, Sezik M, Ozkaya O. Evaluation of a new surgical approach for the treatment of clomiphene citrate-resistant infertility in polycystic ovary syndrome: laparoscopic ovarian multi-needle intervention. Journal of Minimally Invasive Gynecology 2005;12:355-8. 3. Zarei A, Alborzi S, Askary E, Alborzi M, Shahbazi F. Effects of clomiphene citrate for prevention of premature luteinizing hormone surge in those undergoing intrauterine insemination outcome: A randomized, double-blind, placebo-controlled trial. Journal of Advanced Pharmaceutical Technology and Research 2018;9:102-6. 4. Chen J, Feng S, Zeng J, Wu X, Yang M, Tang H et al. Effectiveness of electroacupuncture for polycystic ovary syndrome: study protocol for a randomized controlled trial. Trials 2016;17:256.
No outcomes of interest or outcome data not available: n=4;	1. Mehrabian F, Eessaei F. The laparoscopic ovarian electrocautery versus gonadotropin therapy in infertile women with clomiphene citrate-resistant polycystic ovary syndrome; a randomized controlled trial. JPMA - Journal of the Pakistan Medical Association 2012;62:S42-4. 2. Al-Obaidi MT, Ali ZH, W.I AL-S, E.A.R AL-W, Al-Aubaidy H. Impact of letrozole versus clomiphene citrate on endometrial receptivity in Iraqi women with polycystic ovarian syndrome. Journal of Clinical Pharmacy and Therapeutics 2019. 3. Ashrafinia M, Hosseini R, Moini A, Eslami B, Asgari Z. Comparison of metformin treatment and laparoscopic ovarian diathermy in patients with polycystic ovary syndrome. International Journal of Gynaecology & Obstetrics 2009;107:236-9. 4. Farshchian N, Nezhad ST, Kamangar PB. The combination of letrozole and metformin has a better therapeutic effect on uterine and ovarian arteries in PCOS patients than the combination of clomiphene citrate and metformin. Australasian Medical Journal 2018;11:326-30.
Imbalanced follow-up time: n=2;	1. George SS, George K, Irwin C, Job V, Selvakumar R, Jeyaseelan V et al. Sequential treatment of metformin and clomiphene citrate in clomiphene-resistant women with polycystic ovary syndrome: a randomized, controlled trial. Human Reproduction 2003;18:299-304. 2. Mamonov A, Chaika V. Management of clomiphene resistant patients with PCO syndrome: Metrodin HP versus. laparoscopic electrocoagulation of the ovarian surface (LEOS) 2000:42.
Ongoing trial: n=1	1. NCT03664050. Laparoscopic Ovarian Drilling Versus Letrozole In Clomiphene Citrate Resistant Polycystic Ovary 2018.

Supplemental Table 3 Characteristics of included studies

Study	Country/region	Population	Intervention and comparator	Sample size	Mean age	Mean BMI	Number of cycles	Funding sources
Abd Elgafor 2013	Egypt	CC-resistance	Letrozole+metformin LOD	73 73	24.7 25.1	31.5 32.4	6	not reported
Abdellah 2011	Egypt	CC-resistance	Letrozole LOD	74 73	23.9 23.6	27.3 27.1	6	not reported
Abu Hashim 2010A	Egypt	CC-resistance	Letrozole LOD	128 132	27.3 26.4	26.4 26.6	6	not reported
Abu Hashim 2010B	Egypt	CC-resistance	Letrozole CC+metformin	123 127	28.3 26.2	29.1 30.1	3	not reported
Abu Hashim 2011A	Egypt	CC-failure	CC LOD	89 87	25.2 26.3	25.4 24.7	6	not reported
Abu Hashim 2011B	Egypt	CC-resistance	CC+metformin Gn	75 78	27.5 26.8	26.4 26.3	3	not reported
Abu Hashim 2011C	Egypt	CC-resistance	CC+metformin LOD	138 144	27.2 26.5	26.2 26.1	6	not reported
Bayram 2004	Netherlands	CC-resistance	LOD Gn	83 85	28.5 28.7	27.9 27.3	3*	The Health Insurance Funds Council (OG 97/007), Amstelveen, Netherlands; and Serono Benelux provided financial support for recombinant follicle stimulating hormone during the first eight months of the study when this drug was not funded by the health services.
Begum 2009	Bangladesh	CC-resistance	CC Letrozole	32 32	26.09 25.47	23.632 2.72	6	not reported
Begum 2013	Bangladesh	CC-resistance	CC+metformin Gn+metformin Gn	55 55 55	26.96 26.84 27.15	27.71 28.36 28.98	6	not reported
Davar 2011	Iran	CC-failure	CC+metformin Letrozole+metformin	50 50	29.55 28.54	29.21 28.98	3	Shahid Sadoughi University of Medical Sciences, Yazd, Iran.

De Leo 1999	Italy	CC-failure or CC-resistance	Gn Gn+metformin	10 10	28 29.5	27.7 26.9	2	not reported
Farquhar 2002	New Zealand	CC-resistance	LOD Gn	29 21	29.6 29.6	28.3 27.8	6 3	Auckland Medical Research Foundation, grant 81310
Foroozanfard 2011	Iran	CC-resistance	CC+Gn Letrozole+Gn	60 60	25.33 25.8	24.87 24.12	1	A residency thesis and grant no. 8755 (Kashan University of Medical Sciences)
Ganesh 2009	India	CC-failure or CC-resistance	Letrozole CC+Gn Gn	372 669 346	30.25 30.38 30.82	24.49 24.75 24.08	1	The Council of Scientific and Industrial Research (CSIR)
Ghafarnegad 2010	Iran	CC-failure or CC-resistance	LOD CC+Gn	50 50	26.8 26	28.1 26.72	4	not reported
Ghanem 2013	Egypt	CC-resistance	CC+Gn Gn	87 87	24.8 24.7	33.3 33.2	1	not reported
Hamed 2010	Egypt	CC-resistance	Metformin LOD	55 55	23.6 24.3	35.6 36.1	6	self-funded
Hassan 2017	Egypt	CC-resistance	Letrozole Gn	70 70	28.74 29.95	27.61 27.78	3	not reported
Hwu 2005	Taiwan	CC-resistance	CC CC+metformin	40 40	27.8 29.07	24.11 25.27	1	not reported
Ibrahim 2017	Egypt	CC-resistance	Letrozole LOD	40 40	29.7 28.8	29.21 29.11	6	none
Lazovic 1998	Yugoslavia	CC-resistance	LOD Gn	28 28	NA	NA	6	not reported
Legro 2007	USA	CC-failure	CC+metformin Metformin CC	99 105 106	28.9 29.1 28.8	34.3 35.6 36.4	6	Supported by grants from the National Institutes of Health; Glucophage XR and matching placebo were provided by Bristol-Myers Squibb.
Legro 2014	USA	CC-failure	Letrozole CC	197 167	29.9 30.1	34.2 35.8	5	Supported by grants from the National Institute of Child Health and Human Development; and by the National Center for Research Resources

								and the National Center for Advancing Translational Sciences through an NIH grant to Pennsylvania State University.
Liu 2015	China	CC-resistance	Letrozole LOD	71 70	29.5 28.08	22.5 22.4	6	Shanghai Natural Science Foundation (grant no. 12ZR1434200)
Machado 2012	Brazil	CC-resistance	CC CC+metformin	15 21	27.1 27.7	28.3 30.6	6	not reoprted
Malkawi 2002	Jordan	CC-failure or CC-resistance	CC CC+metformin	12 16	29 29	27.8 27.5	6	not reported
Ng 2001	Hong Kong	CC-resistance	CC+metformin CC	10 10	30.5 32	24.1 23.8	4	The Committee on Reasearch and Conference Grants, The University of Hong Kong
Palomba 2004	Italy	CC-resistance	Metformin LOD	60 60	26.8 27.5	28.1 27.6	6	not reported
Palomba 2005	Italy	CC-failure or CC-resistance	Gn Gn+metformin	35 35	26.9 26.2	26.4 26.5	6	not reported
Palomba 2010	Italy	CC-resistance	LOD CC+metformin	25 25	28.2 27.5	29.8 30.2	6	no outside funding
Rezk 2018	Egypt	CC-resistance	CC+metformin Letrozole	105 104	24.6 24.2	24.2 23.7	3	not reported
Safdarian 2012	Iran	CC-resistance	Letrozole Letrozole+Gn	26 33	26.15 27.76	26.72 26.76	1	not reported
Seyedoshohadaei 2016	Iran	CC-failure	CC Letrozole	50 50	30.04 29.62	NA	NA	Kurdistan University of Medical Sciences
Sharma 2010	India	CC-resistance	Gn CC+Gn Letrozole+Gn	185 181 178	NA	NA	>1	not reported
Sohrabvand 2006	Iran	CC-failure	CC+metformin Letrozole+metformin	30 30	29.55 28.24	30.21 29.98	>1	not reported
Sturrock 2002	UK	CC-resistance	CC CC+metformin	14 12	NA	NA	6	not reported
Tasdemir 2004	Turkey	CC-resistance	Gn Gn+metformin	16 16	30.6 31.8	29 28.5	1	not reported

van Santbrink 2005	Netherlands	CC-resistance	Gn Gn+metformin	9 11	28* 28*	34* 38*	2	not reported
Vandermolen 2001	USA	CC-resistance	CC CC+metformin	15 12	30 29	35.4 37.6	8	National Institute of Child Health and Human Development, National Institutes of Health
Vegetti 1998	Italy	CC-resistance	Gn LOD	13 16	NA	NA	6	not reported
Weiss 2017	Netherlands	CC-failure	CC Gn	335 331	29.9 29.7	25.2 25.5	6	The Netherlands Organization for Health Research and Development
Yadav 2017	India	CC-resistance	LOD Gn	45 44	26.11 26.23	25 24.94	3*	not reported
Yarali 2002	Turkey	CC-resistance	Gn Gn+metformin	16 16	28.4 29.7	29.6 28.6	2	not reported

Footnote: * refers to the number of cycles included in the analysis; # refers to median; NA: not available; Gn: Gonadotrophins; LOD: laparoscopic ovarian drilling

Supplemental Table 4 Pairwise meta-analyses

Comparison (versus CC)	Live birth/ongoing pregnancy Odds ratio (95% confidence interval)	Multiple pregnancy Odds ratio (95% confidence interval)	Clinical pregnancy Odds ratio (95% confidence interval)	Ovulation Odds ratio (95% confidence interval)	Miscarriage Odds ratio (95% confidence interval)
CC+metformin	1.84 (0.72-4.73)	3.21 (0.13-77.89)	1.97 (1.12-3.47)	1.95 (1.19-3.17)	6.15 (1.11-34.09)
LOD	1.15 (0.73-1.80)	0.11 (0.01-2.08)	1.16 (0.78-1.72)	NA	1.23 (0.39-3.87)
Letrozole	1.62 (1.10-2.38)	0.57 (0.10-3.34)	1.30 (1.12-1.51)	1.29 (1.08-1.56)	2.34 (0.26-21.10)
Gonadotrophins	1.22 (1.04-1.45)	0.89 (0.32-2.41)	1.67 (1.25-2.23)	NA	2.21 (1.10-4.43)
CC+gonadotrophins	NA	NA	NA	NA	NA
Gonadotrophins+metformin	NA	NA	NA	NA	NA
Letrozole+metformin	NA	NA	NA	NA	NA
Letrozole+gonadotrophins	NA	NA	NA	NA	NA

Footnote: NA: not available

Supplemental Table 5 Subgroup analysis

Comparison (versus CC) live birth/ongoing pregnancy	CC-failure Odds ratio (95% confidence interval)	CC-resistance Odds ratio (95% confidence interval)
CC+metformin	1.09 (0.56-2.13)	5.23 (1.17-23.44)
LOD	1.21 (0.64-2.32)	6.80 (1.23-37.76)
Letrozole	1.82 (1.14-2.92)	9.40 (1.62-54.40)
Gonadotrophins	1.45 (1.07-1.97)	12.87 (2.32-71.45)
CC+Gonadotrophins	1.59 (0.47-5.36)	15.59 (2.45-99.37)
Gonadotrophins+metformin	2.63 (0.96-7.23)	38.29 (6.57-223.29)
Letrozole+metformin	4.91 (1.03-23.44)	NA

Footnote: NA: not available

Supplemental Table 6 Sensitivity analysis on RCTs at low risk of bias at randomisation

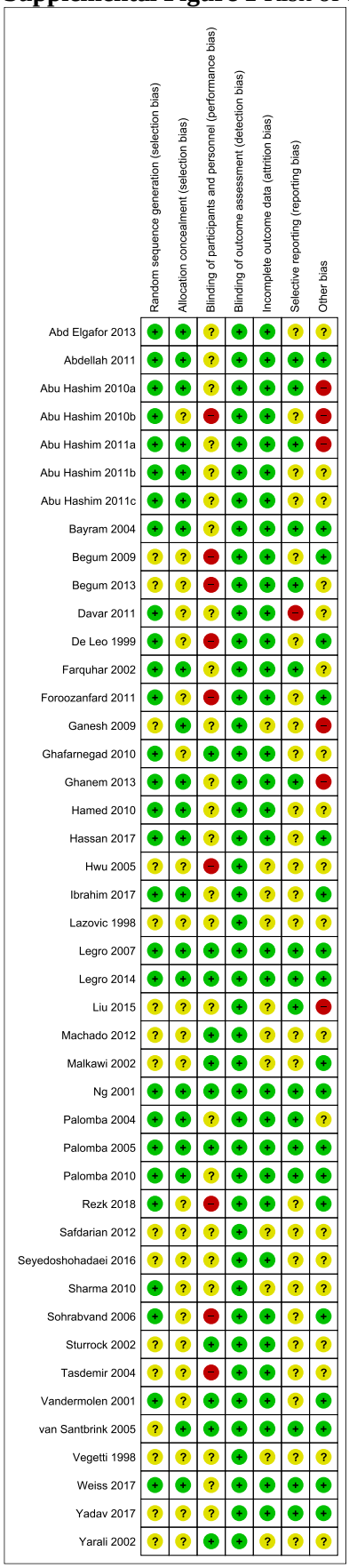
Comparison (versus CC) live birth/ongoing pregnancy	**Sensitivity analysis Odds ratio (95% confidence interval)**
CC+metformin	1.11 (0.63-1.95)
LOD	1.18 (0.82-1.71)
Letrozole	1.60 (1.07-2.40)
Gonadotrophins	1.57 (1.18-2.08)
CC+Gonadotrophins	1.90 (0.89-4.04)
Gonadotrophins+metformin	2.84 (1.04-7.73)

Supplemental Table 7 Network analyses for the other outcomes

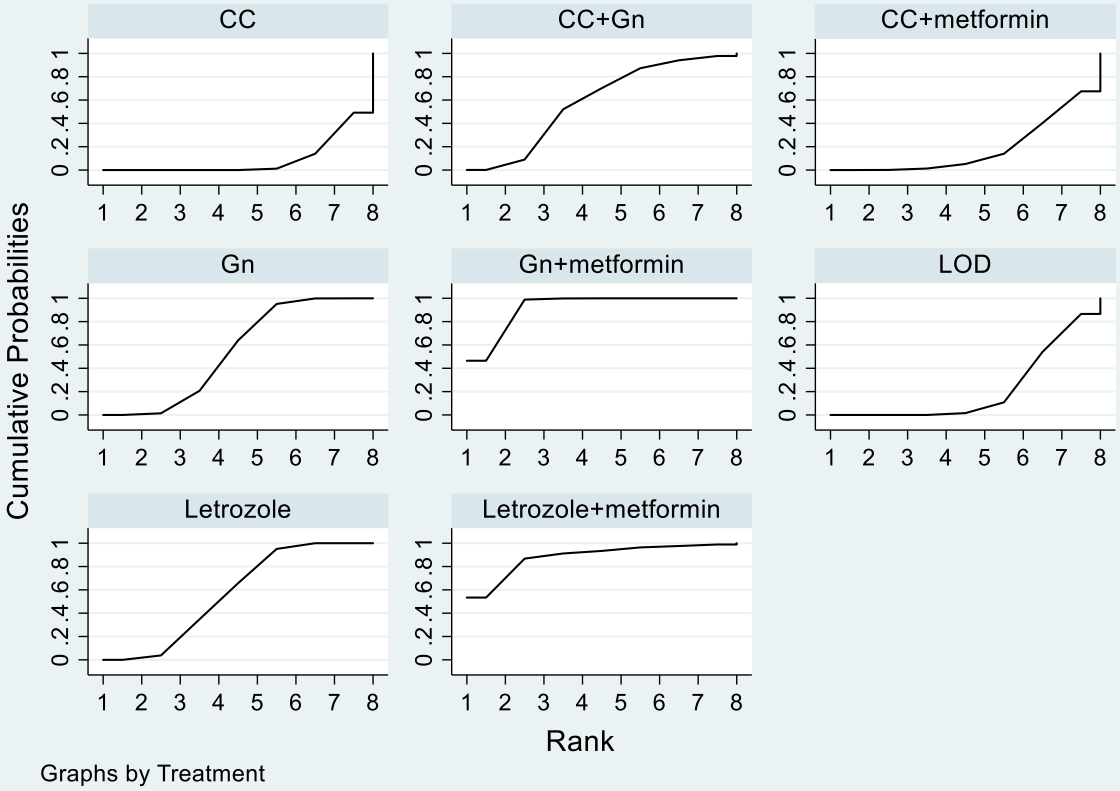
Comparison (versus CC)	**Multiple pregnancy Odds ratio (95% confidence interval) 21 RCTs 3072 women**	**Clinical pregnancy Odds ratio (95% confidence interval) 20 RCTs 3944 women**	**Ovulation Odds ratio (95% confidence interval) 41 RCTs 6836 women**	**Miscarriage Odds ratio (95% confidence interval) 33 RCTs 6362 women**
CC+metformin	0.44 (0.13-1.55)	1.33 (0.90-1.95)	2.44 (1.38-4.34)	1.90 (0.91-3.97)
LOD	0.12 (0.03-0.51)	1.51 (1.02-2.23)	2.52 (1.14-5.56)	1.56 (0.82-2.97)
Letrozole	0.59 (0.17-2.08)	2.30 (1.61-3.29)	5.82 (3.06-11.03)	1.66 (0.94-2.94)
Gonadotrophins	1.05 (0.43-2.57)	2.19 (1.48-3.24)	13.96 (6.64-29.37)	2.22 (1.28-3.84)
CC+gonadotrophins	0.57 (0.12-2.58)	1.68 (1.01-2.80)	2.04 (0.71-5.89)	1.76 (0.84-3.68)
Gonadotrophins+metformin	0.58 (0.12-2.79)	5.77 (2.94-11.31)	21.89 (6.77-70.81)	3.07 (0.88-10.68)
Letrozole+metformin	NA	2.22 (1.04-4.74)	2.66 (0.66-10.70)	1.29 (0.32-5.26)
Letrozole+gonadotrophins	0.17 (0.02-1.64)	2.16 (1.17-4.00)	NA	1.92 (0.62-5.88)

Footnote: NA: not available

Supplemental Figure 1 Risk of bias of included studies

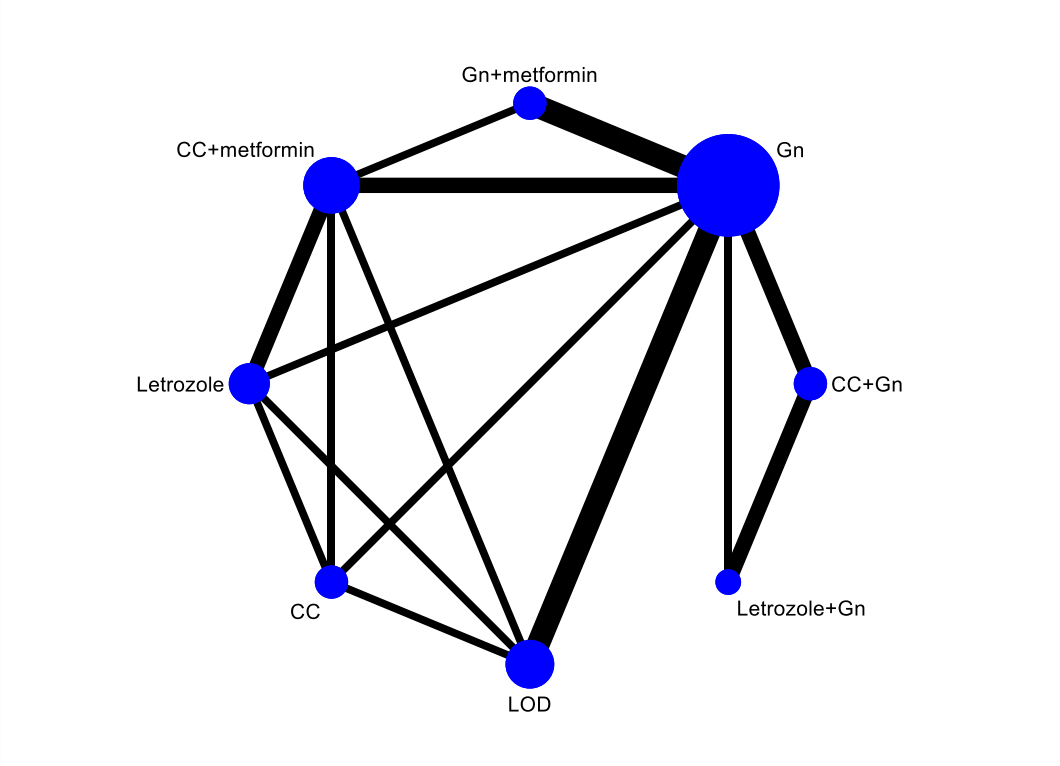


Supplemental Figure 2 Cumulative rankograms of interventions for live birth/ongoing pregnancy

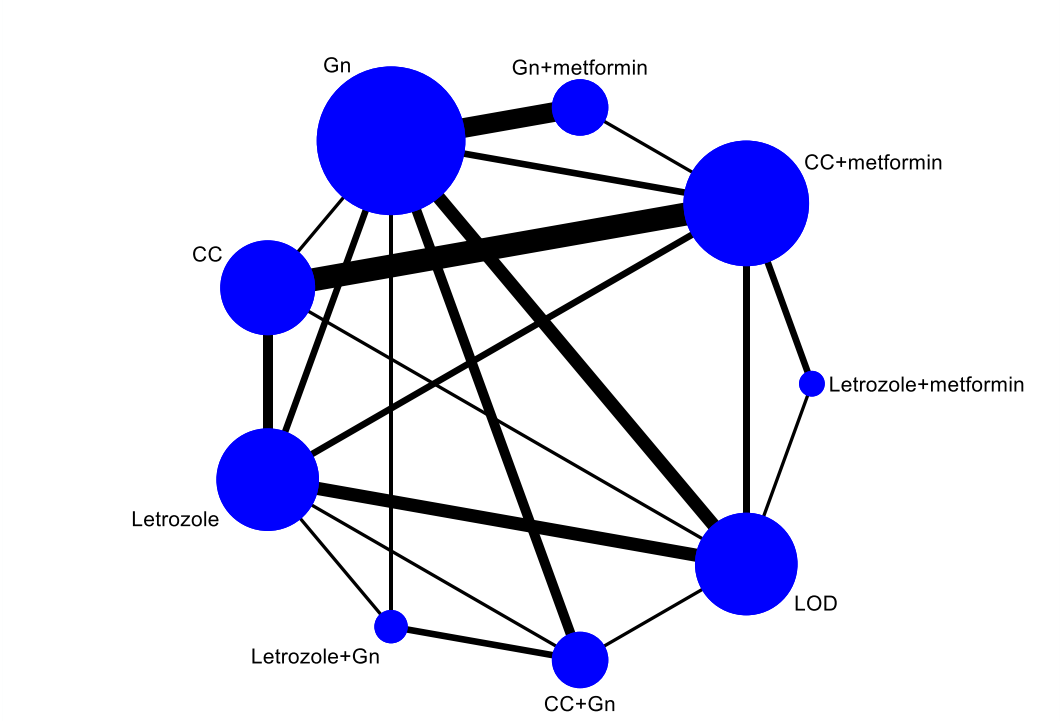


Each cumulative rankogram illustrates the cumulative probability of each ranking (from the best to the worst rank) for each intervention in terms of live birth/ongoing pregnancy.

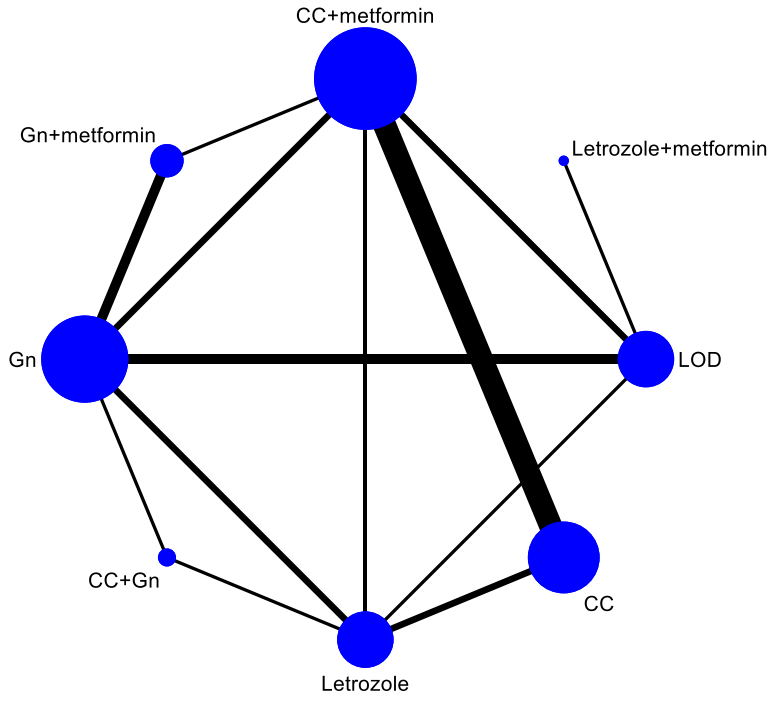
Supplemental Figure 3 Network plot for multiple pregnancy



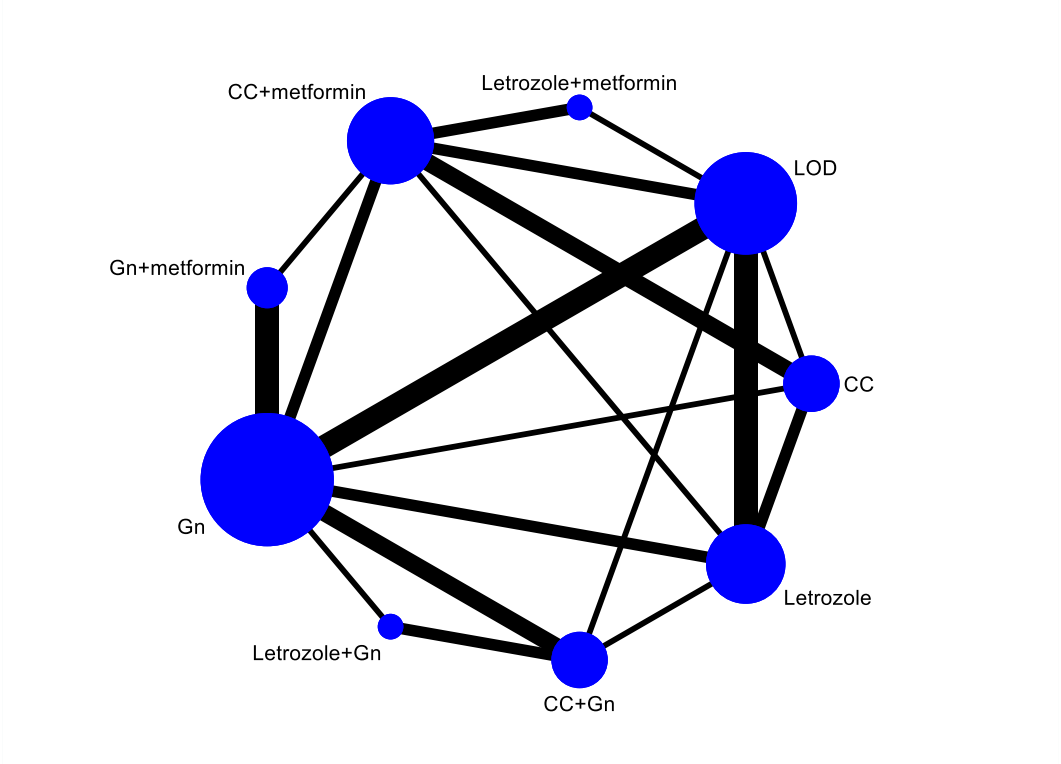
Supplemental Figure 4 Network plot for clinical pregnancy



Supplemental Figure 5 Network plot for ovulation



Supplemental Figure 6 Network plot for miscarriage



Supplemental materials for Chapter 5

Appendix 1 Cochrane Gynaecology and Fertility Group (CGF) search strategy
Searched 6 September 2018
Procite platform
Keywords CONTAINS "unexplained and endometriosis related infertility" or "unexplained infertility" or "unexplained subfertility" or "idiopathic infertility "or "idiopathic male infertility" or "idiopathic subfertility" or Title CONTAINS "unexplained and endometriosis related infertility" or "unexplained infertility" or "unexplained subfertility" or "idiopathic infertility" or "idiopathic male infertility" or "idiopathic subfertility" (374 hits)

Appendix 2 Cochrane Central Register of Studies Online (CRSO) search strategy
Searched 6 September 2018
CRSO web platform
#1 MESH DESCRIPTOR Infertility EXPLODE ALL TREES 2759
#2 unexplained:TI,AB,KY 1712
#3 idiopathic:TI,AB,KY 7295
#4 #2 OR #3 8953
#5 #1 AND #4 373
#6 (unexplain* adj5 infertil*):TI,AB,KY 483
#7 (unexplain* adj5 subfertil*):TI,AB,KY 74
#8 (idiopathic adj5 subfertil*):TI,AB,KY 11
#9 (idiopathic adj5 infertil*):TI,AB,KY 94
#10 (unknown adj5 subfertil*):TI,AB,KY 1
#11 (unknown adj5 infertil*):TI,AB,KY 1
#12 #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 720

Appendix 3 MEDLINE search strategy
Searched from 1946 to 6 September 2018
Ovid platform
1 exp Infertility/ and unexplained.tw. (1901)
2 exp Infertility/ and idiopathic.tw. (1700)
3 (unexplain* adj5 infertil*).tw. (2090)
4 (unexplain* adj5 subfertil*).tw. (157)
5 (idiopathic adj5 subfertil*).tw. (74)
6 (idiopathic adj5 infertil*).tw. (1222)
7 (unknown adj3 infertil*).tw. (170)
8 (unknown adj3 subfertil*).tw. (11)
9 (unexplained adj3 steril*).tw. (56)
10 (idiopathic adj3 steril*).tw. (54)
11 (unknown adj3 steril*).tw. (48)
12 or/1-11 (4512)
13 exp Clomiphene/ (5115)
14 clomifene.tw. (127)
15 clomiphene.tw. (4875)
16 Serophene.tw. (4)
17 clomid.tw. (176)
18 selective estrogen receptor modulators/ or exp raloxifene hydrochloride/ or exp tamoxifen/ (21795)
19 selective estrogen receptor modulator*.tw. (2803)
20 (SERMs or SERM).tw. (2009)
21 (raloxifene or tamoxifen).tw. (23603)
22 or/13-21 (36987)
23 Aromatase Inhibitors/ (5733)
24 Aromatase inhibitor*.tw. (6687)
25 letrozole.tw. (2481)
26 (femara or anastrozole).tw. (1675)

27 (anti-?estrogen* or anti?estrogen*).tw. (8947)
28 or/23-27 (17912)
29 exp follicle stimulating hormone/ or exp follicle stimulating hormone, beta subunit/ or exp glycoprotein hormones, alpha subunit/ or exp menotropins/ or exp urofollitropin/ (38849)
30 Follicle Stimulating Hormone*.tw. (18222)
31 (FSH or rFSH or recFSH).tw. (33058)
32 (uFSH or rhFSH).tw. (233)
33 (hpFSH or pFSH).tw. (203)
34 (follitropin or Gonal F).tw. (705)
35 (menotropin* or menopur).tw. (207)
36 corifollitropin.tw. (90)
37 (urofollitropin or pergonal or bravelle* or follitrin).tw. (206)
38 Follistim*.tw. (12)
39 (Puregon or humegon or menogon).tw. (89)
40 human menopausal gonadotrop?in.tw. (1783)
41 growth hormone.tw. (53592)
42 HMG.tw. (13823)
43 gonadotrop?in*.tw. (60770)
44 or/29-43 (157278)
45 expectant management.tw. (2298)
46 watchful waiting.tw. (2284)
47 (watch and wait).tw. (750)
48 Coitus/ (7072)
49 coitus.tw. (2693)
50 intercourse.tw. (18110)
51 sex*.tw. (651487)
52 or/45-51 (662646)
53 exp Insemination, Artificial/ (11188)
54 intrauterine insemination*.tw. (2295)
55 artificial insemination*.tw. (6200)
56 superovulat*.tw. (3265)
57 IUI.tw. (1587)
58 or/53-56 (17342)
59 exp embryo transfer/ or exp fertilization in vitro/ or exp sperm injections, intracytoplasmic/ (38494)
60 embryo transfer*.tw. (10716)
61 vitro fertili?ation.tw. (21146)
62 ivf.tw. (21404)
63 icsi.tw. (7513)
64 intracytoplasmic sperm injection*.tw. (6494)
65 (blastocyst adj2 transfer*).tw. (877)
66 exp reproductive techniques, assisted/ or exp insemination, artificial/ or exp ovulation induction/ (63849)
67 assisted reproduct*.tw. (13076)
68 ovulation induc*.tw. (3941)
69 (ovar* adj2 stimulat*).tw. (6529)
70 ovarian hyperstimulation.tw. (4741)
71 COH.tw. (1579)
72 (ovar* adj2 induc*).tw. (3910)
73 (modified adj3 cycle*).tw. (560)
74 (natural adj3 cycle*).tw. (2396)
75 MNC IVF.tw. (23)
76 (NCIVF or NC-IVF).tw. (18)
77 unstimulated ivf.tw. (18)
78 (unstimulated adj2 in vitro fertili?ation).tw. (13)
79 (artificial adj3 cycle$).tw. (449)
80 or/59-79 (87813)
81 22 or 28 or 44 or 52 or 58 or 80 (914386)
82 12 and 81 (2487)

83 randomized controlled trial.pt. (467907)
84 controlled clinical trial.pt. (92625)
85 randomized.ab. (421185)
86 randomised.ab. (84107)
87 placebo.tw. (196867)
88 clinical trials as topic.sh. (184705)
89 randomly.ab. (296832)
90 trial.ti. (187190)
91 (crossover or cross-over or cross over).tw. (77604)
92 or/83-91 (1229120)
93 exp animals/ not humans.sh. (4493841)
94 92 not 93 (1131490)
95 82 and 94 (493)

Appendix 4 Embase search strategy
Searched from 1980 to 6 September 2018
Ovid platform
1 (exp infertility/ or exp infertility therapy/) and unexplained.tw. (3790)
2 (exp infertility/ or exp infertility therapy/) and idiopathic.tw. (3240)
3 (unexplain* adj5 infertil*).tw. (3122)
4 (unexplain* adj5 subfertil*).tw. (252)
5 (idiopathic adj5 subfertil*).tw. (89)
6 (idiopathic adj5 infertil*).tw. (1739)
7 (unknown adj3 infertil*).tw. (262)
8 (unknown adj3 subfertil*).tw. (14)
9 (unexplained adj3 steril*).tw. (59)
10 (idiopathic adj3 steril*).tw. (60)
11 (unknown adj3 steril*).tw. (56)
12 or/1-11 (7404)
13 exp clomifene/ (4436)
14 clomifene.tw. (215)
15 clomiphene.tw. (5229)
16 Serophene.tw. (194)
17 clomid.tw. (922)
18 exp selective estrogen receptor modulator/ (7325)
19 exp raloxifene/ (10783)
20 exp tamoxifen citrate/ or exp tamoxifen/ (58156)
21 selective estrogen receptor modulator*.tw. (3748)
22 (SERMs or SERM).tw. (2979)
23 (raloxifene or tamoxifen).tw. (33459)
24 or/13-23 (78982)
25 exp aromatase inhibitor/ (28231)
26 Aromatase inhibitor*.tw. (10361)
27 letrozole.tw. (4470)
28 (femara or anastrozole).tw. (3652)
29 (anti-?estrogen* or anti?estrogen*).tw. (10420)
30 or/25-29 (38435)
31 exp follitropin/ (48940)
32 exp human menopausal gonadotropin/ (8642)
33 exp urofollitropin/ (1649)
34 Follicle Stimulating Hormone*.tw. (18482)
35 (FSH or rFSH or recFSH).tw. (39571)
36 (uFSH or rhFSH).tw. (334)
37 (hpFSH or pFSH).tw. (207)
38 (follitropin or Gonal F).tw. (2940)
39 (menotropin* or menopur).tw. (773)
40 corifollitropin.tw. (190)
41 (urofollitropin or pergonal or bravelle* or follitrin).tw. (2034)

42 Follistim*.tw. (268)
43 (Puregon or humegon or menogon).tw. (2081)
44 human menopausal gonadotrop?in.tw. (1863)
45 growth hormone.tw. (55310)
46 HMG.tw. (17377)
47 gonadotrop?in*.tw. (61170)
48 or/31-47 (175256)
49 expectant management.tw. (3317)
50 watchful waiting.tw. (3340)
51 (watch and wait).tw. (1386)
52 exp coitus/ (5008)
53 coitus.tw. (2579)
54 intercourse.tw. (22847)
55 sex*.tw. (810058)
56 or/49-55 (823896)
57 exp artificial insemination/ (15778)
58 intrauterine insemination*.tw. (3376)
59 artificial insemination*.tw. (5478)
60 superovulat*.tw. (3537)
61 IUI.tw. (2883)
62 or/49-61 (843368)
63 exp fertilization in vitro/ (60536)
64 exp embryo transfer/ (27677)
65 exp intracytoplasmic sperm injection/ (18393)
66 embryo transfer*.tw. (16874)
67 vitro fertili?ation.tw. (27109)
68 ivf.tw. (35769)
69 icsi.tw. (14240)
70 intracytoplasmic sperm injection*.tw. (8545)
71 (blastocyst adj2 transfer*).tw. (1989)
72 exp infertility therapy/ (87213)
73 exp artificial insemination/ (15778)
74 exp ovulation induction/ (13068)
75 assisted reproduct*.tw. (19632)
76 ovulation induc*.tw. (5192)
77 (ovar* adj2 stimulat*).tw. (9965)
78 ovarian hyperstimulation.tw. (6858)
79 COH.tw. (2177)
80 (ovar* adj2 induc*).tw. (4609)
81 (modified adj3 cycle*).tw. (775)
82 (natural adj3 cycle*).tw. (3197)
83 MNC IVF.tw. (37)
84 (NCIVF or NC-IVF).tw. (47)
85 unstimulated ivf.tw. (30)
86 (unstimulated adj2 in vitro fertili?ation).tw. (18)
87 (artificial adj3 cycle$).tw. (528)
88 or/63-87 (122326)
89 24 or 30 or 48 or 56 or 62 or 88 (1159178)
90 Clinical Trial/ (939390)
91 Randomized Controlled Trial/ (506064)
92 exp randomization/ (79110)
93 Single Blind Procedure/ (32096)
94 Double Blind Procedure/ (148976)
95 Crossover Procedure/ (56068)
96 Placebo/ (307810)
97 Randomi?ed controlled trial$.tw. (183941)
98 Rct.tw. (29057)
99 random allocation.tw. (1783)
100 randomly.tw. (381272)

101 randomly allocated.tw. (30135)
102 allocated randomly.tw. (2330)
103 (allocated adj2 random).tw. (792)
104 Single blind$.tw. (21113)
105 Double blind$.tw. (182169)
106 ((treble or triple) adj blind$).tw. (803)
107 placebo$.tw. (269794)
108 prospective study/ (463707)
109 or/90-108 (2118634)
110 case study/ (55585)
111 case report.tw. (348767)
112 abstract report/ or letter/ (1017866)
113 or/110-112 (1413483)
114 109 not 113 (2069749)
115 (exp animal/ or animal.hw. or nonhuman/) not (exp human/ or human cell/ or (human or humans).ti.) (5490558)
116 114 not 115 (1926969)
117 12 and 89 and 116 (1070)

Appendix 5 PsycINFO search strategy
Searched from 1806 to 6 September 2018
Ovid platform
1 exp INFERTILITY/ and unexplained.tw. (40)
2 exp INFERTILITY/ and idiopathic.tw. (16)
3 (unexplain* adj5 infertil*).tw. (36)
4 (unexplain* adj5 subfertil*).tw. (2)
5 (idiopathic adj5 infertil*).tw. (18)
6 (unknown adj3 infertil*).tw. (10)
7 (unexplained adj3 steril*).tw. (1)
8 (idiopathic adj3 steril*).tw. (2)
9 (unknown adj3 steril*).tw. (2)
10 or/1-9 (71)
11 random*.ti,ab,hw,id. (181184)
12 trial*.ti,ab,hw,id. (166702)
13 controlled stud*.ti,ab,hw,id. (11453)
14 placebo*.ti,ab,hw,id. (38171)
15 ((singl* or doubl* or trebl* or tripl*) and (blind* or mask*)).ti,ab,hw,id. (27288)
16 (cross over or crossover or factorial* or latin square).ti,ab,hw,id. (27952)
17 (assign* or allocat* or volunteer*).ti,ab,hw,id. (152430)
18 treatment effectiveness evaluation/ (22271)
19 mental health program evaluation/ (2045)
20 exp experimental design/ (54262)
21 or/11-20 (480042)
22 10 and 21 (6)

Appendix 6 CINAHL search strategy
Searched from 1961 to 6 September 2018
Ebsco platform

#	Query	Results
S23	S10 AND S22	102
S22	S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21	1,255,308
S21	TX allocat* random*	9,041
S20	(MH "Quantitative Studies")	20,295
S19	(MH "Placebos")	10,838
S18	TX placebo*	52,082
S17	TX random* allocat*	9,041

S16	(MH "Random Assignment")	50,544
S15	TX randomi* control* trial*	153,119
S14	TX ((singl* n1 blind*) or (singl* n1 mask*)) or TX ((doubl* n1 blind*) or (doubl* n1 mask*)) or TX ((tripl* n1 blind*) or (tripl* n1 mask*)) or TX ((trebl* n1 blind*) or (trebl* n1 mask*))	972,401
S13	TX clinic* n1 trial*	227,640
S12	PT Clinical trial	86,040
S11	(MH "Clinical Trials+")	244,190
S10	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9	309
S9	TX(idiopathic N3 steril*)	2
S8	TX(unknown N3 subfertil*)	1
S7	TX(unknown N3 infertil*)	19
S6	TX(idiopathic N5 infertil*)	60
S5	TX(idiopathic N5 subfertil*)	5
S4	TX(unexplain* N5 subfertil*)	30
S3	TX (unexplain* N5 infertil*)	185
S2	(MM "Infertility") and TX idiopathic	64
S1	(MM "Infertility") and TX unexplained	147

Appendix 7 Data and analyses

7.1 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]

7.2 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]

7.3 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]

Appendix 8 Summary of findings tables

8.1 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	166 per 1000	**235 per 1000** (117 to 416)	**OR 1.55** (0.67 to 3.58)	⊕⊕⊖⊖ LOW[a]

	(no direct evidence available; only indirect evidence used here)				
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.

8.2 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]
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.

8.3 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]
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.

8.4 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

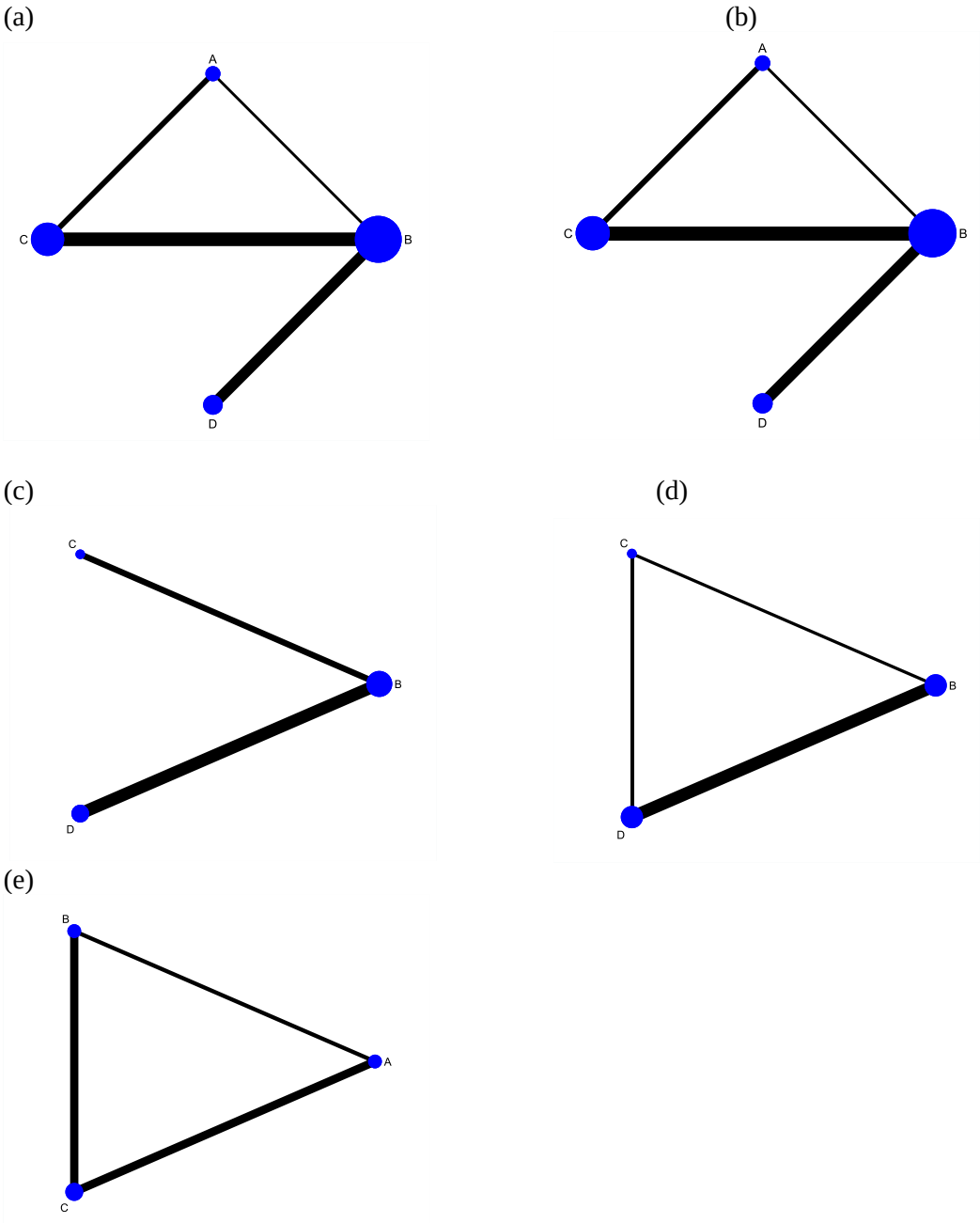
Supplemental materials for Chapter 6

Table S1 Results of global inconsistency assessment

Outcome	Chi^2	P
Clinical pregnancy at the longest follow-up time	16.62	0.001
Clinical pregnancy within 6 months	1.80	0.18
Clinical pregnancy within 3 months	1.38	0.24
Clinical pregnancy within 9 months	-*	-
Clinical pregnancy within 12 months	0.26	0.61
Ongoing pregnancy within 3 months	1.50	0.22
Live birth resulting from pregnancy within 6 months	0.28	0.60
Miscarriage within 6 months	0.85	0.36

Consistency refers to agreement between direct evidence and indirect evidence. Differences in study populations, interventions (comparators), outcomes and risks of bias can result in inconsistency. Global consistency was tested by using the 'design-by-interaction' chi square test. A P-value < 0.1 indicates evidence of global inconsistency in the network. There was no evidence of global inconsistency in all the following networks of different outcomes.*There was no closed loop in this network and therefore global inconsistency was not tested.

Figure S1 Network plots for clinical pregnancy within 3 months (a), 6 months (b), 9 months (c) and 12 months (d), and live birth resulting from pregnancy within 6 months (e) after intervention.



These network plots illustrate available head-to-head (direct) comparisons between different interventions. If there is a line between two interventions, direct comparison is available between this comparison. The width of the lines is proportional to the number of trials comparing each pair of treatments, and the size of each node is proportional to the number of studies investigating the respective interventions. A: no tubal flushing; B: water-based contrast; C: oil-based contrast; D: water- and oil-based contrast.

Figure S2 Risk of bias contributions

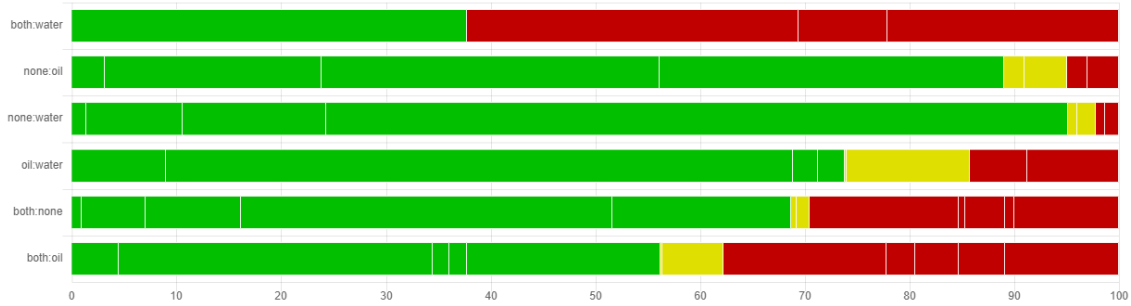
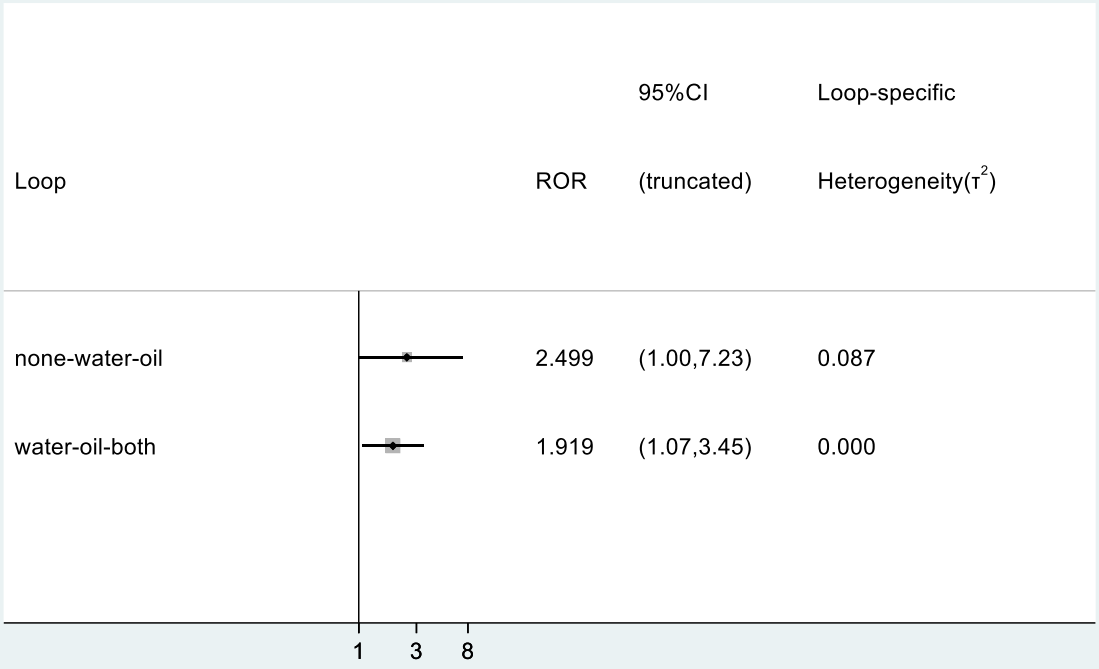


Figure shows proportions (%) of different quality of evidence in each comparison for the main outcome (clinical pregnancy within 6 months). Green – low risk of bias; yellow – moderate risk of bias; red – high risk of bias.

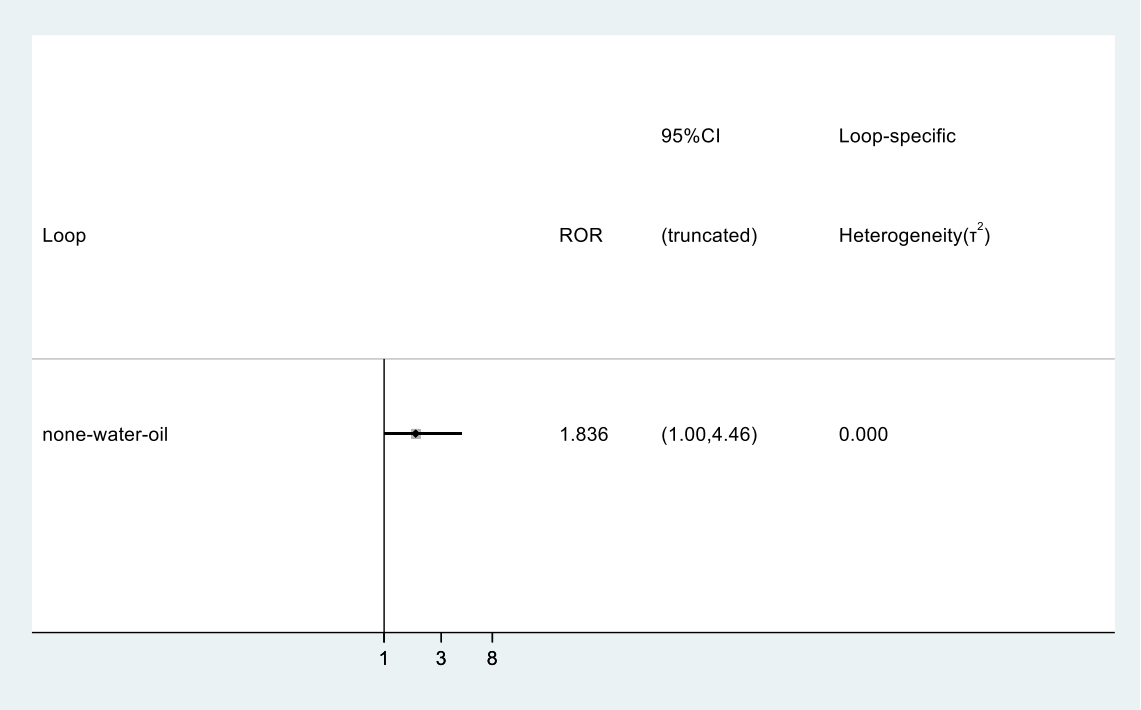
Figure S3 Inconsistency plots for clinical pregnancy at the longest follow-up time (a) clinical pregnancy within 6 months (b), 3 months (c), 12 months (d) and live birth within 6 months after intervention.

(a) Inconsistency plot for clinical pregnancy at the longest follow-up time

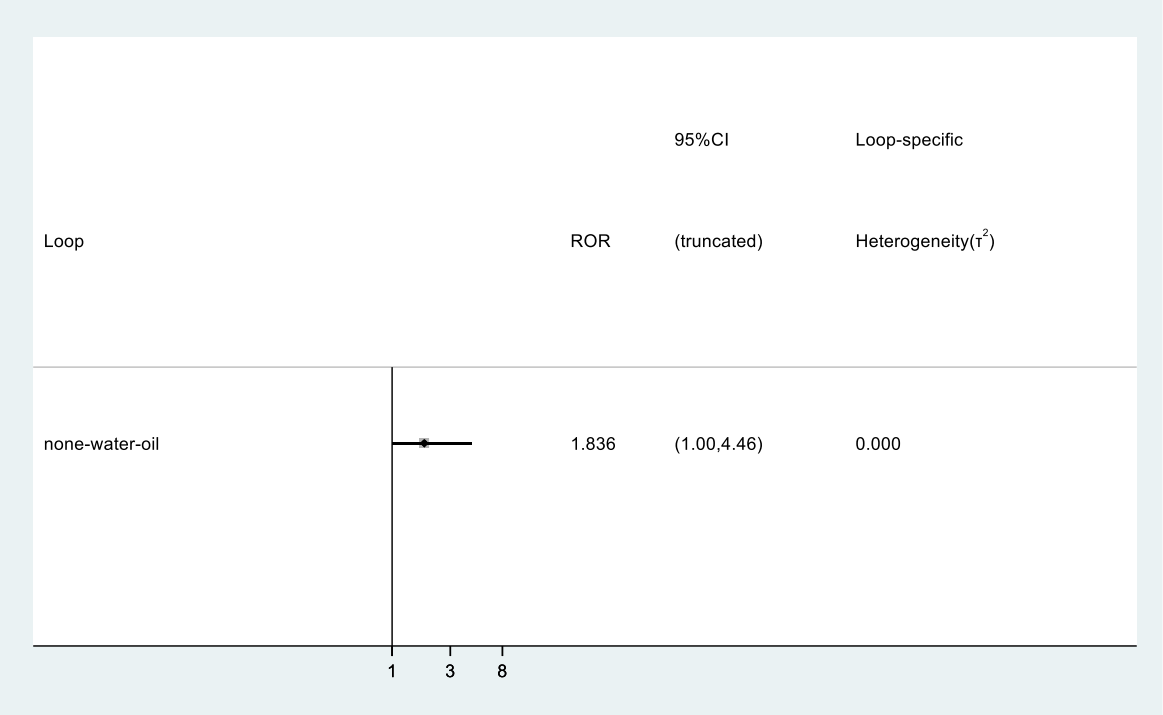
The following inconsistency plot shows that there are two closed loops in the network: none-water-oil loop and water-oil-both loop. These are statistically significant inconsistency in both loops (none-water-oil: p = 0.091; water-oil-both: p = 0.029).



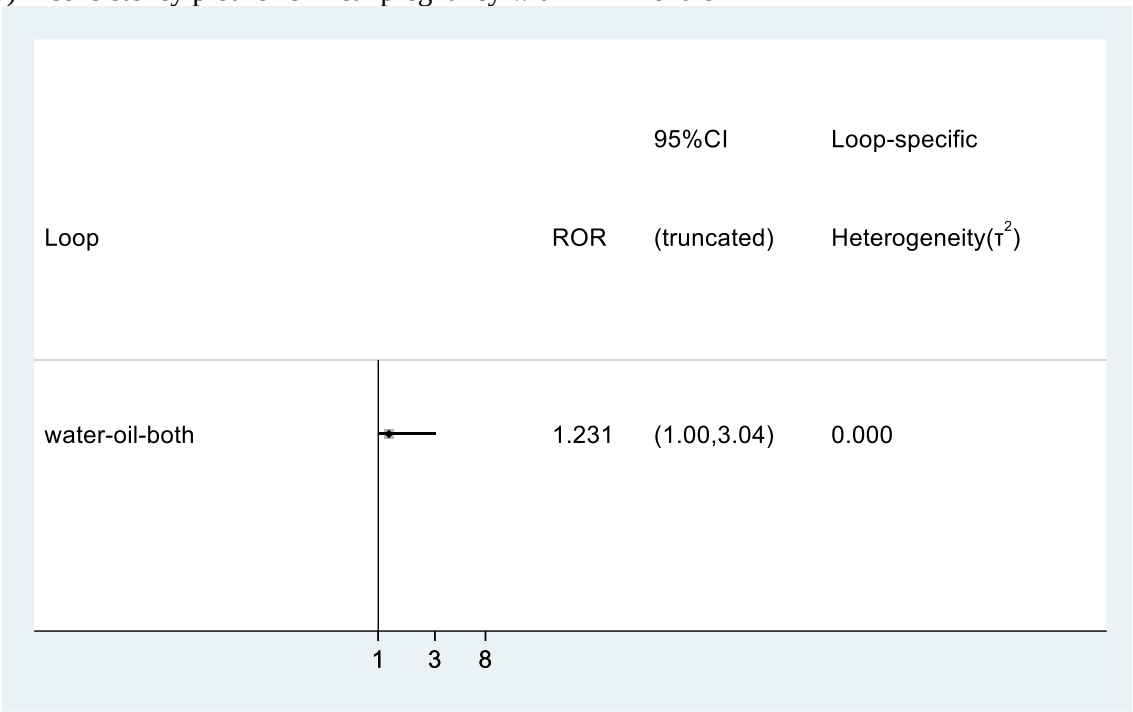
(b) Inconsistency plot for clinical pregnancy within 6 months



(c) Inconsistency plot for clinical pregnancy within 3 months



(d) Inconsistency plot for clinical pregnancy within 12 months



(e) Inconsistency plot for live birth within 6 months

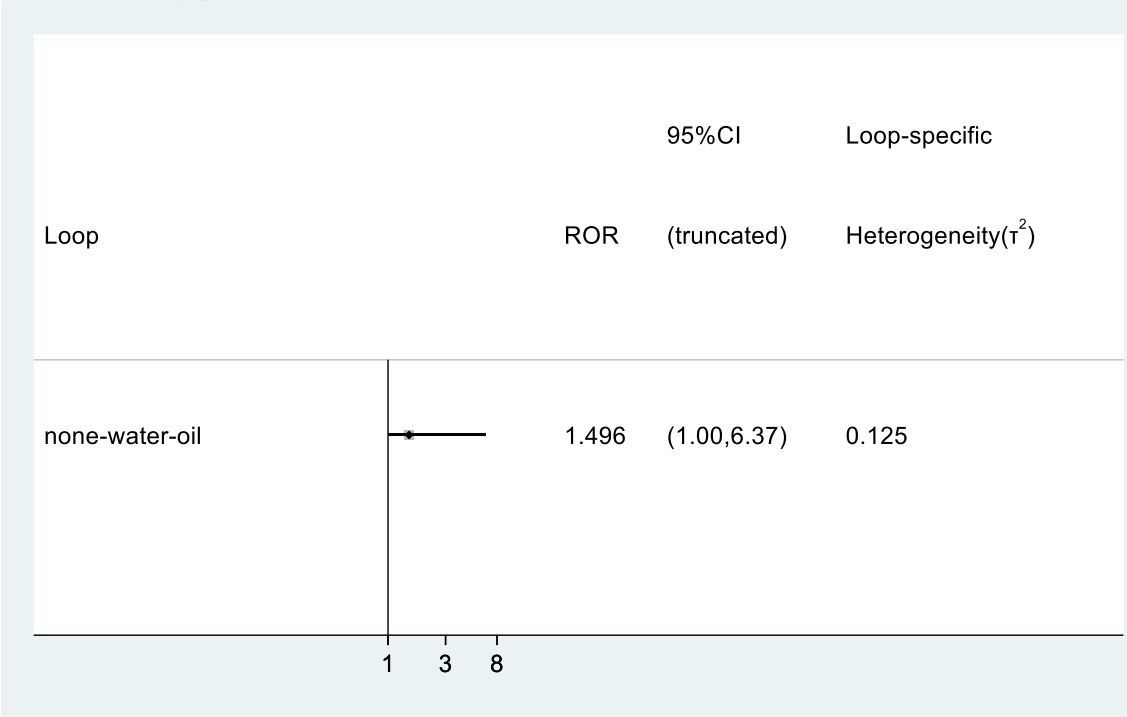
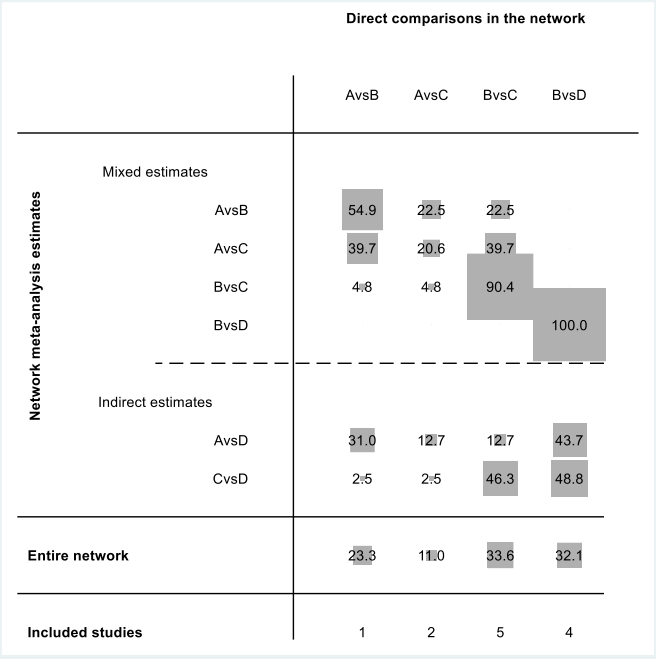


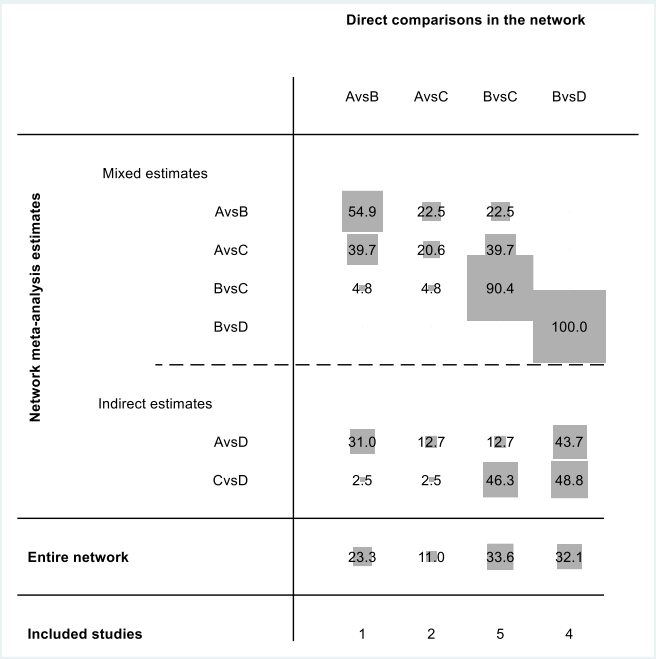
Figure S4 Contribution plots for clinical pregnancy within 6 months (a), 3 months (b), 9 months (c) and 12 months (d), live birth resulting from pregnancy within 6 months (e) and miscarriage within 6 months (f) after intervention.

(a) Contribution plot for clinical pregnancy within 6 months

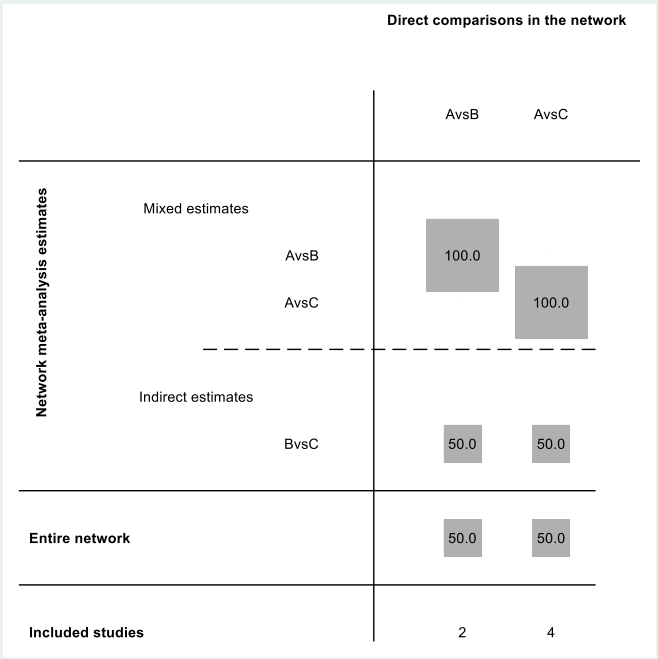
There are four direct comparisons in the network: A vs B, A vs C, B vs C and B vs D. The contribution of these direct comparisons to the network are 23.5%, 11.0%, 33.6% and 32.1%, respectively. (A: no tubal flushing; B: water-based contrast; C: oil-based contrast; D: water- and oil-based contrast)



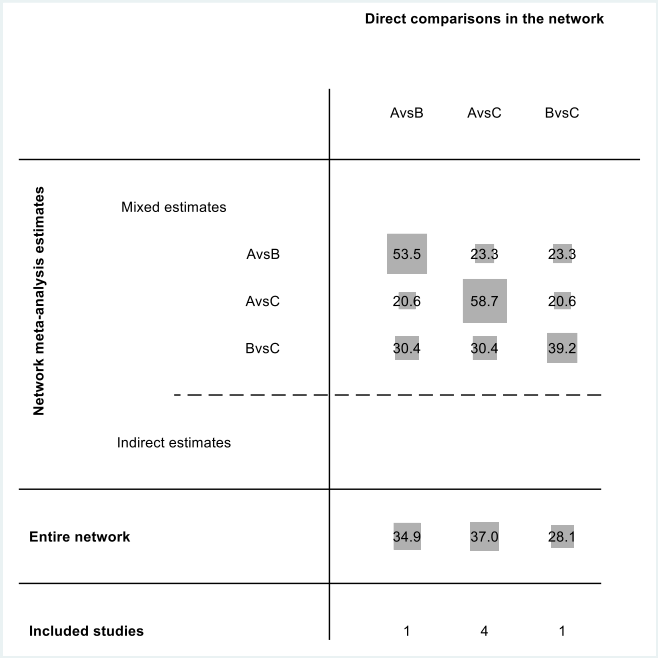
(b) Contribution plot for clinical pregnancy within 3 months (A: no tubal flushing; B: water-based contrast; C: oil-based contrast; D: water- and oil-based contrast)



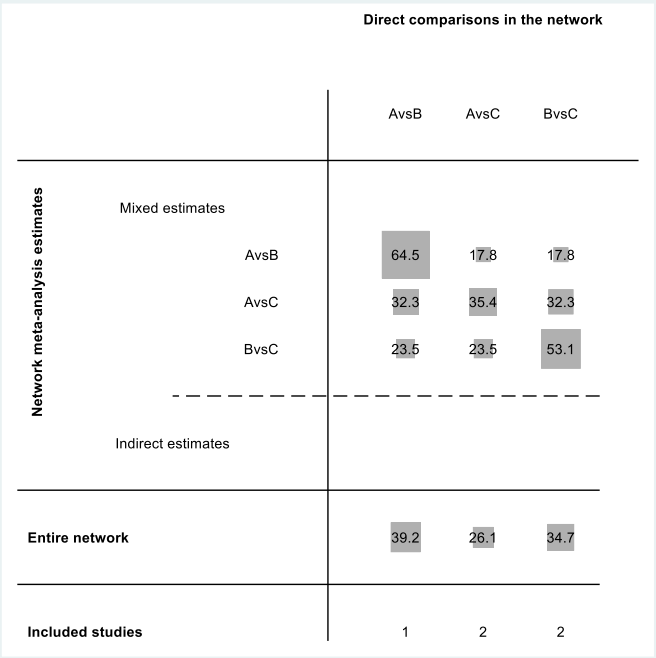
(c) Contribution plot for clinical pregnancy within 9 months (A: water-based contrast; B: oil-based contrast; C: water- and oil-based contrast)



(d) Contribution plot for clinical pregnancy within 12 months (A: water-based contrast; B: oil-based contrast; C: water- and oil-based contrast)



(e) Contribution plot for live birth resulting from pregnancy within 6 months (A: water-based contrast; B: oil-based contrast; C: water- and oil-based contrast)



(f) Contribution plot for miscarriage within 6 months (A: water-based contrast; B: oil-based contrast; C: water- and oil-based contrast)

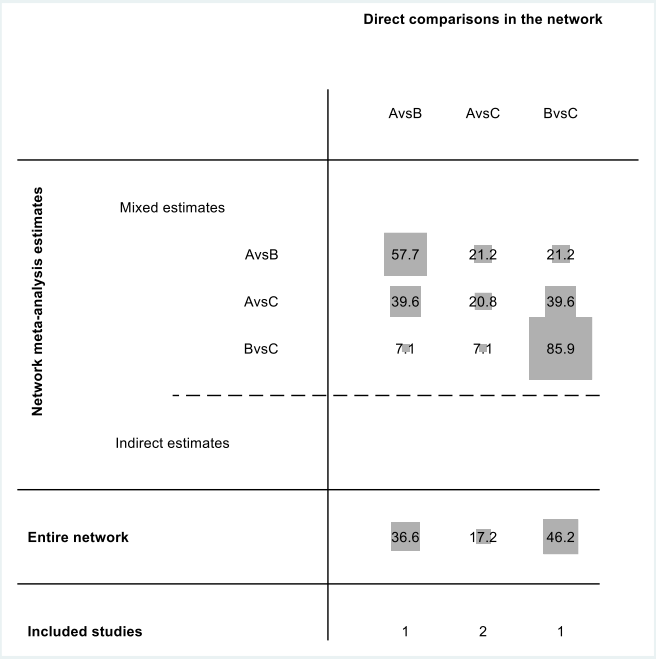
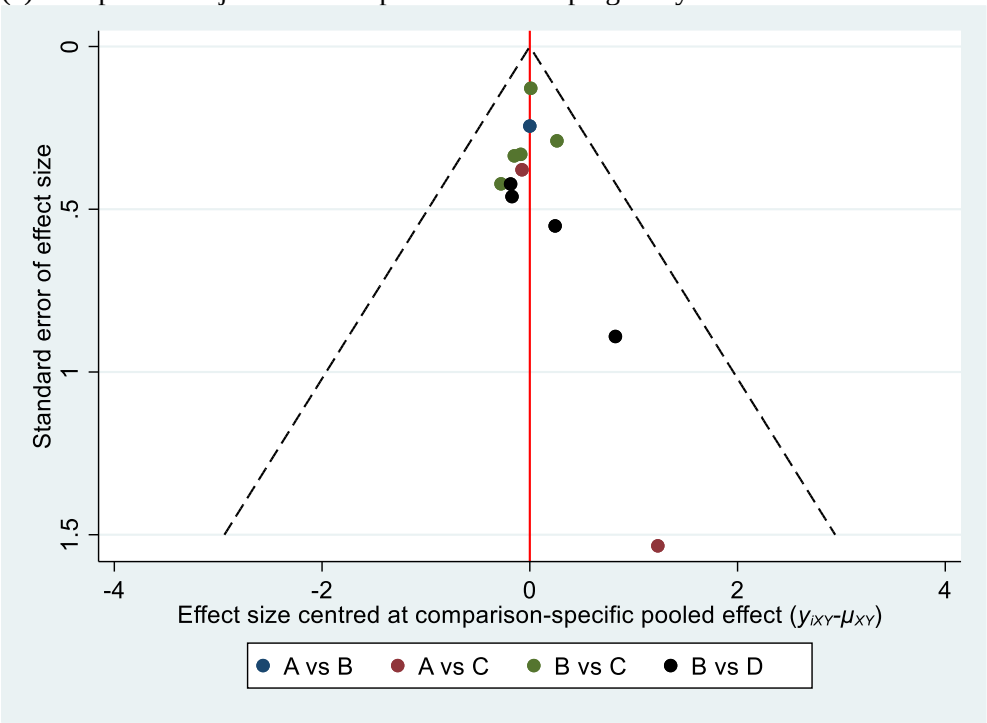
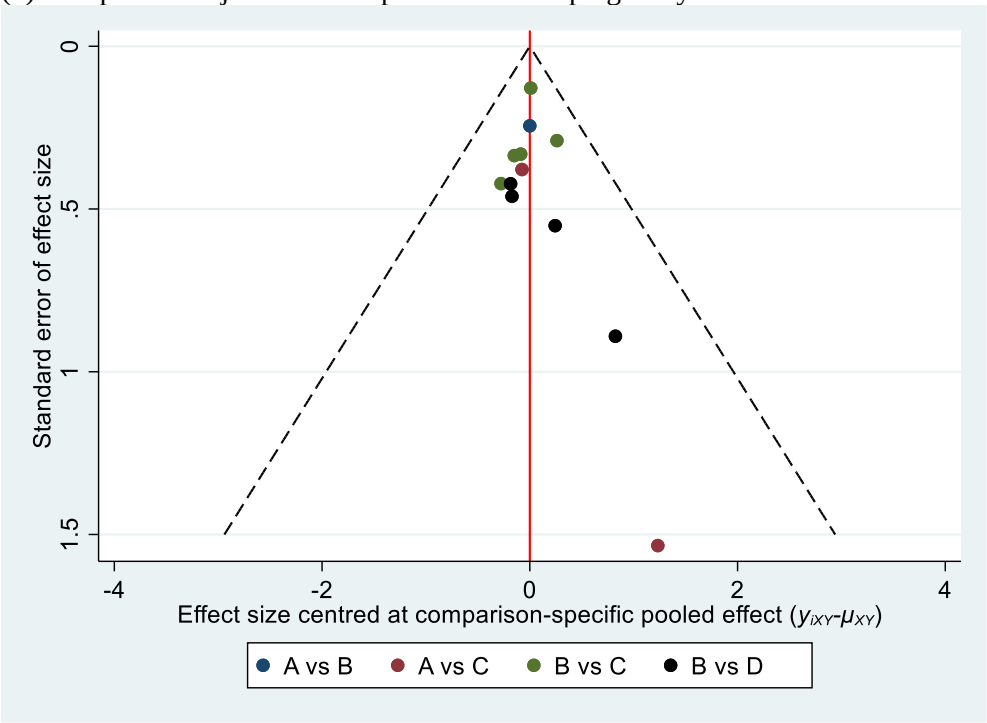


Figure S5 Comparison adjusted funnel plots for clinical pregnancy within 6 months (a) and 3 months (b) after intervention.

(a) Comparison adjusted funnel plot for clinical pregnancy within 6 months



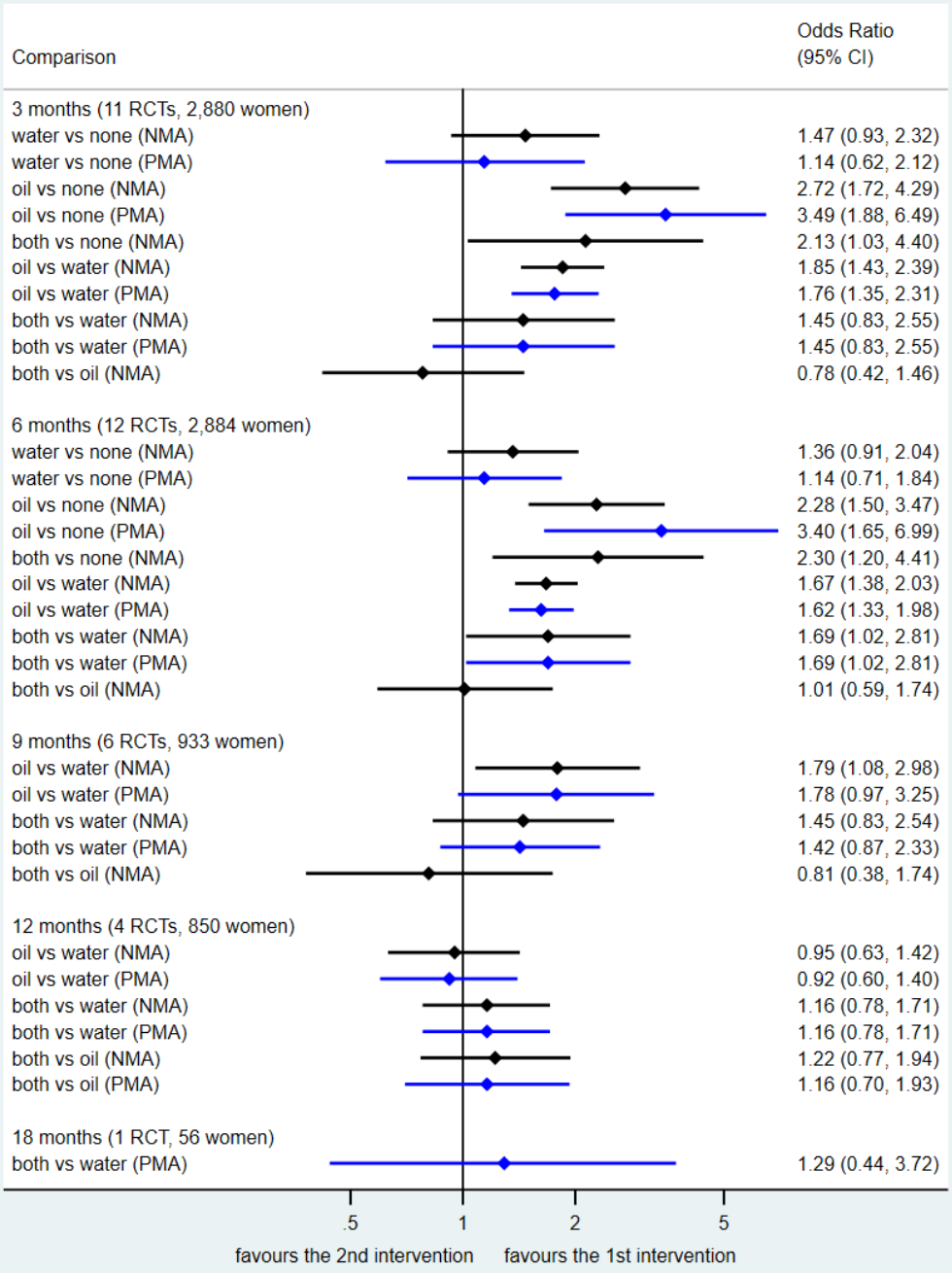
(b) Comparison adjusted funnel plot for clinical pregnancy within 3 months



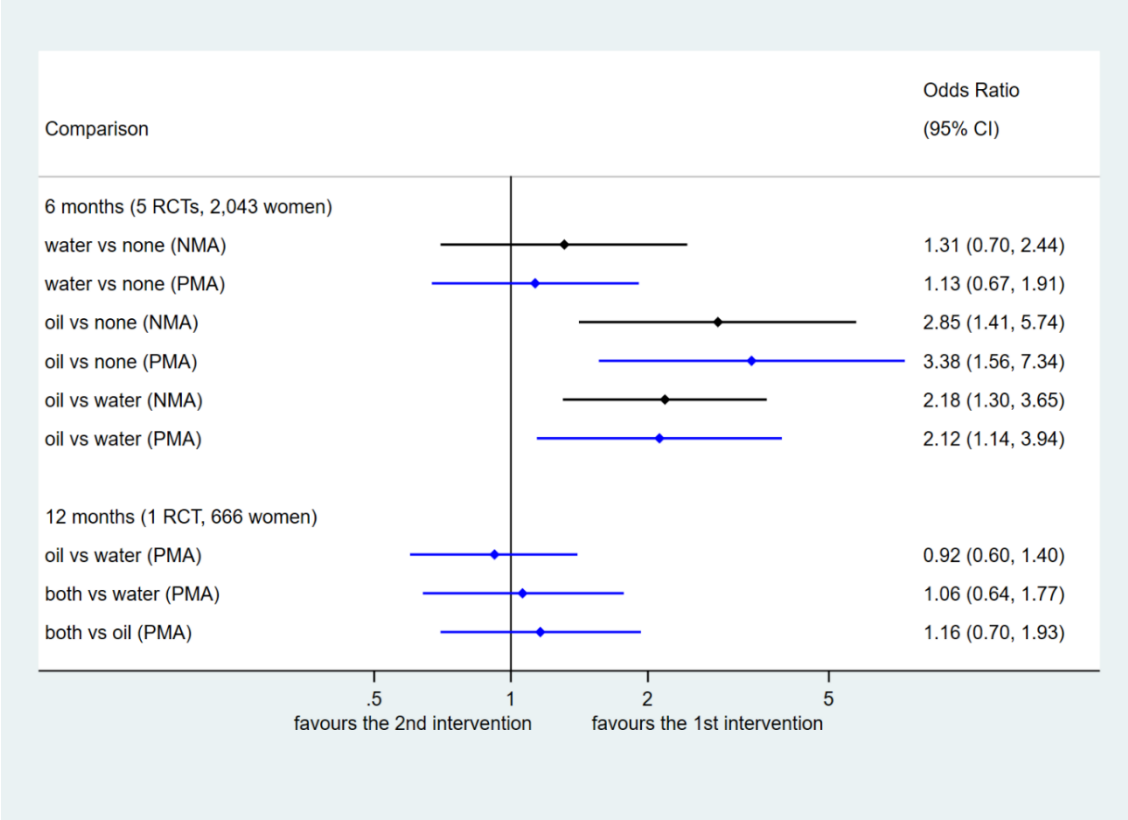
The red line represents the null hypothesis that the study-specific effect sizes do not differ from the respective comparison-specific pooled effect estimates. Different colours refer to different comparisons. (A: no tubal flushing; B: water-based contrast; C: oil-based contrast; D: water- and oil-based contrast)

Figure S6 Network and pairwise meta-analyses for clinical pregnancy at different timepoints (a), live birth at different timepoints (b), ongoing pregnancy resulting from pregnancy within 6 months after intervention (c), miscarriage (d) and ectopic pregnancy and adverse events (e).
(a) Network and pairwise meta-analyses for clinical pregnancy at different timepoints

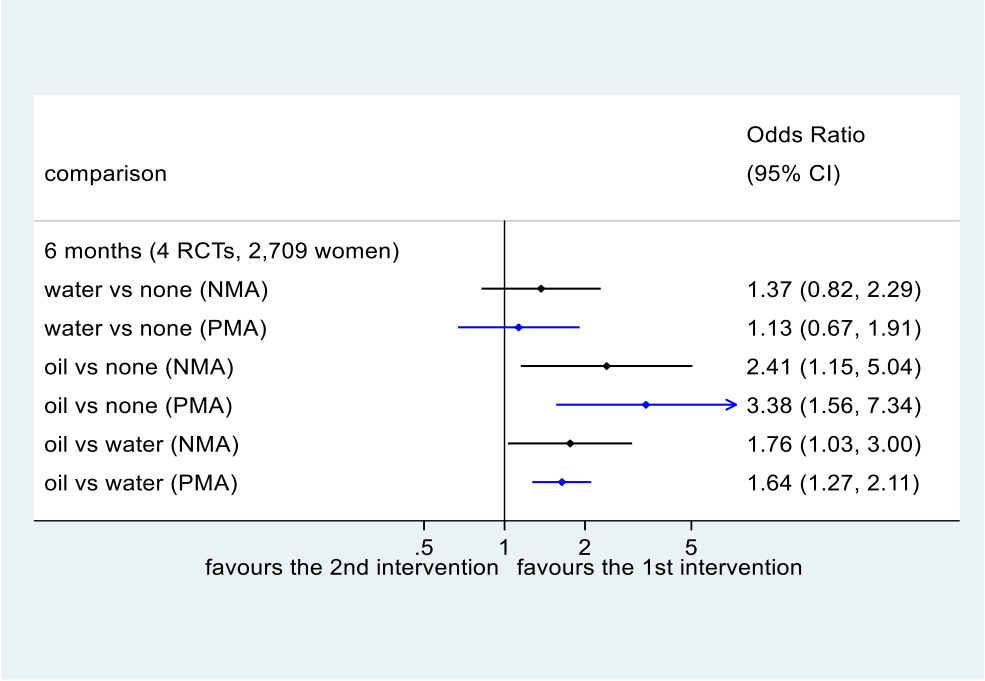
Clinical pregnancies at different time points are pooled in subgroups. Both network meta-analysis and pairwise meta-analysis are included in this forest plot. Diamonds and lines represent odds ratios (ORs) and relevant 95% confidence intervals (CIs), respectively. ORs and 95% CIs of network meta-analyses are illustrated in black while those of pairwise meta-analyses are illustrated in blue.



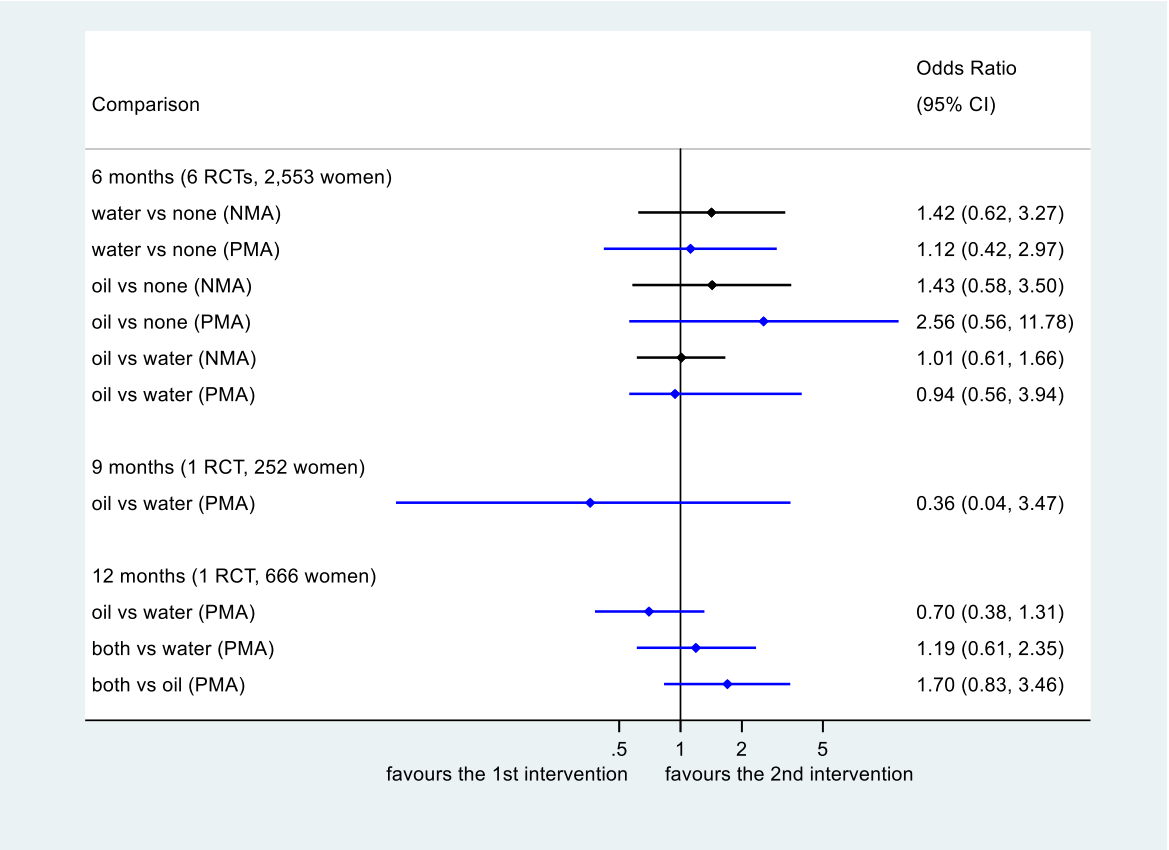
(b) Network and pairwise meta-analyses for live birth at different timepoints



(c) Network and pairwise meta-analyses for ongoing pregnancy resulting from pregnancy within 6 months



(d) Network and pairwise meta-analyses for miscarriage

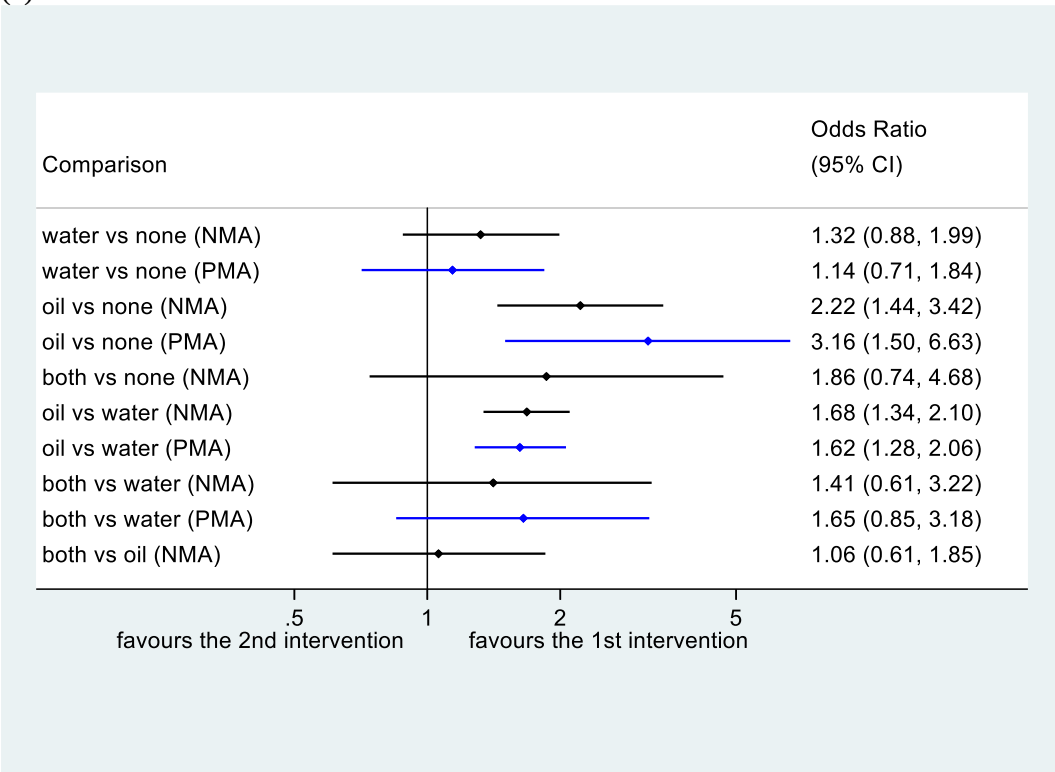


(e) Pairwise meta-analyses for ectopic pregnancy and adverse events

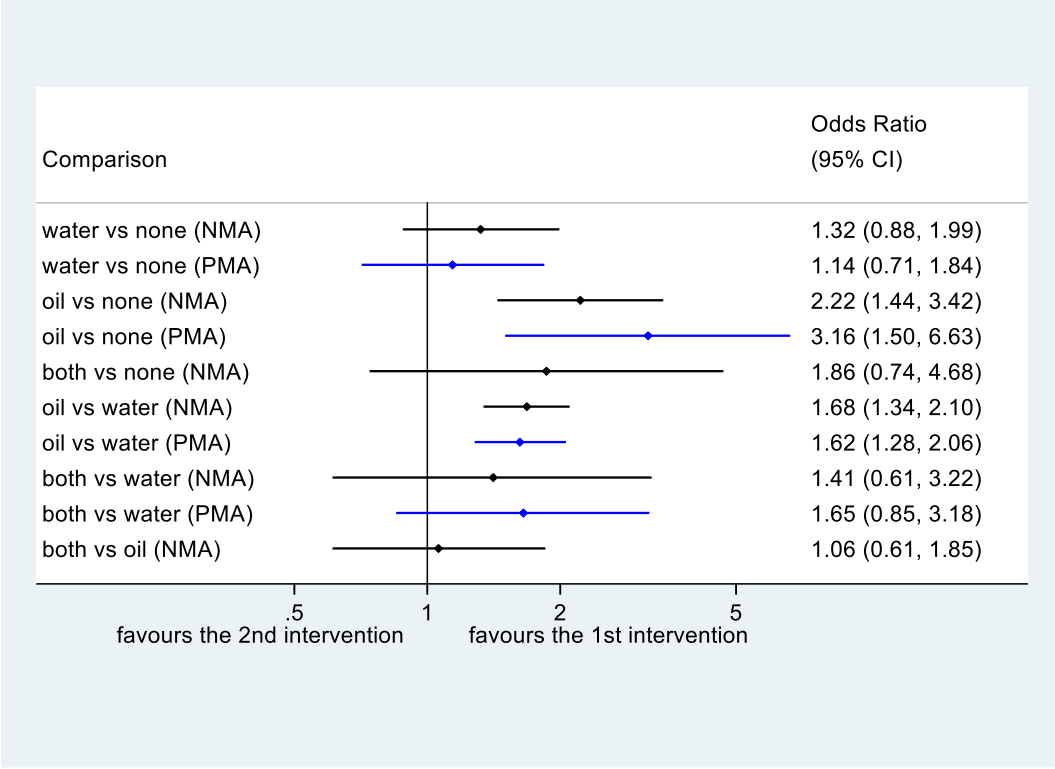
Comparisons	No. of studies	No. of participants	OR (95% CI)	I^2
Ectopic pregnancy within 6 months				
water vs none	1	334	0.99 (0.06, 15.93)	-
oil vs none	1	158	3.54 (0.14, 88.18)	-
oil vs water	2	1250	0.667 (0.13, 3.48)	0
Ectopic pregnancy within 9 months				
oil vs water	1	242	3.02 (0.12, 74.99)	-
Ectopic pregnancy within 12 months				
oil vs water	1	533	0.47 (0.09, 2.60)	-
both vs water	1	303	0.48 (0.05, 4.38)	-
both vs oil	1	406	1.03 (0.09, 11.42)	-
Pelvic infection				
oil vs water	2	662	0.23 (0.04, 1.27)	0
Intravasation				
oil vs water	3	793	5.06 (2.29, 11.18)	0

Figure S7 Sensitivity analysis for clinical pregnancy within 6 months when including only studies with low risk of bias (a) and when excluding participants with missing outcome data (b).

(a)



(b)



Supplemental materials for Chapter 8

Supplementary Table 1 List of included studies without IPD and reasons

List of included studies without IPD	Reasons
Aygen 2007; Badawy 2009; Badawy 2011; Fleming 2002; Keikha 2011; Khorram 2006; Mobusher 2014; Santonocito 2009; Tang 2006*; Zain 2009	Data loss (n = 10)
Moussa 2016	Legal reasons (n = 1)
Abuelghar 2013; Atay 2006; Ayaz 2013; Banerjee Ray 2012; Basirat 2012; Boostanfar 2001; Chen 2016; Dasari 2009; Dehbashi 2009; Hossein-Rashidi 2016; Jahan 2015; Karimzadeh 2007; Karimzadeh 2010; Lopez 2004; Lorzadeh 2011; Maged 2015; Robinson 2003; Roy 2012; Selim 2012; Seyedoshohadaei 2012; Sharief 2015; Sheikh-El-Arab Elsedeek 2011; Zeinalzadeh 2010	No response (n = 23)
Beigi 2006; Boudhraa 2010; Cudmore 1966; El-Biely 2001; Garcia 1985; Johnson 1966	IPD not sought due to insufficient contact information (n = 6)
Fatima 2018; Topçu 2017	IPD not sought as studies were identified after the data requesting timeline (n = 2)

*Note: Although IPD of baseline and other outcomes in this study were provided, IPD of outcomes of interest for this IPD meta-analysis were not available.

Supplementary Table 2 Sensitivity analyses for live birth

Comparison	Sensitivity analyses	Number of RCTs	Number of participants	Risk Ratio (RR)	95% confidence interval (CI)	I^2
Letrozole vs CC	RCTs with low risk of bias at allocation concealment	2	909	1.42	1.14-1.76	0
	Treatment naïve women with PCOS	3	627	1.41	1.11-1.79	0
CC+metformin vs CC	RCTs with low risk of bias at allocation concealment	3	714	1.02	0.76-1.37	33.2%
	Treatment naïve women with PCOS	5	662	1.06	0.83-1.34	3.9%

Supplementary Table 3 IPD availability bias

Comparison (vs. CC)	Network meta-analyses of RCTs with and without IPD RR (95% CI)	Network meta-analyses of RCTs with IPD RR (95% CI)	Network meta-analyses of RCTs without IPD RR (95% CI)
Live birth	**27 RCTs 5257 women**	**12 RCTs 3437 women**	**15 RCTs 1820 women**
Placebo	0.58 (0.31-1.07)	0.56 (0.26-1.20)	NA
Metformin	0.90 (0.64-1.28)	0.87 (0.51-1.47)	0.95 (0.57-1.58)
CC + Metformin	1.27 (0.91-1.78)	1.18 (0.73-1.90)	1.70 (0.88-3.30)
Letrozole	1.46 (1.09-1.95)	1.42 (0.79-2.55)	1.47 (1.07-2.02)
Tamoxifen	1.16 (0.61-2.18)	NA	1.12 (0.65-1.94)
Gonadotrophins	1.31 (0.73-2.34)	1.22 (0.45-3.34)	1.45 (0.65-3.22)
Clinical pregnancy	**62 RCTs 9356 women**	**20 RCTs 3962 women**	**42 RCTs 5394 women**
Placebo	0.49 (0.33-0.71)	0.61 (0.37-1.01)	0.30 (0.16-0.57)
Metformin	1.06 (0.83-1.34)	0.94 (0.67-1.34)	1.13 (0.80-1.59)
CC + Metformin	1.46 (1.21-1.76)	1.34 (1.02-1.76)	1.62 (1.23-2.13)
Letrozole	1.37 (1.16-1.61)	1.48 (1.07-2.05)	1.30 (1.08-1.58)
Tamoxifen	0.91 (0.66-1.25)	0.72 (0.26-1.95)	0.91 (0.65-1.26)
Gonadotrophins	1.34 (0.87-2.08)	1.22 (0.64-2.31)	1.57 (0.81-3.06)

This table shows the results of network meta-analyses of RCTs with IPD and network meta-analyses of all eligible RCTs on live birth and clinical pregnancy. The results are presented in the comparisons of different interventions versus CC for live birth and clinical pregnancy, respectively.

Supplementary Table 4 List of investigators of the primary RCTs

Primary RCTs	Investigators
CLET trial (Amer 2017)	S.A. Amer, J. Smith, A. Mahran, and P. Fox, A. Fakis
Bayar 2006	Ülkü Bayar, Mustafa Basaran, Sibel Kiran, Ayhan Coskun and Sener Gezer
COFFI trial (Homburg 2012)	R. Homburg, M.L. Hendriks, T.E. König, R.A. Anderson, A.H. Balen, M. Brincat, T. Child, M. Davies, T. D'Hooghe, A. Martinez, M. Rajkhowa, R. Rueda-Saenz, P. Hompes and C.B. Lambalk
PCOSMIC trial* (Johnson 2010)	N.P. Johnson, A.W. Stewart, J. Falkiner, C.M. Farquhar, S. Milsom, V.-P. Singh, Q.L. Okonkwo, K.L. Buckingham, REACT-NZ (REproductionAnd Collaborative Trials in New Zealand)
Kar 2012	Sujata Kar
Kar 2015	Sujata Kar and Smriti Sanchita
Leanza 2014	V Leanza, L Coco, F Grasso, G Leanza, G Zarbo, and M Palumbo.
PPCOS I trial (Legro 2007)	Richard S. Legro, Huiman X. Barnhart, William D. Schlaff, Bruce R. Carr, Michael P. Diamond, Sandra A. Carson, Michael P. Steinkampf, Christos Coutifaris, Peter G. McGovern, Nicholas A. Cataldo, Gabriella G. Gosman, John E. Nestler, Linda C. Giudice, Phyllis C. Leppert, and Evan R. Myers, for the Cooperative Multicenter Reproductive Medicine Network
PPCOS II trial (Legro 2014)	Richard S. Legro, Robert G. Brzyski, Michael P. Diamond, Christos Coutifaris, William D. Schlaff, Peter Casson, Gregory M. Christman, Hao Huang, Qingshang Yan, Ruben Alvero, Daniel J. Haisenleder, Kurt T. Barnhart, G. Wright Bates, Rebecca Usadi, Scott Lucidi, Valerie Baker, J.C. Trussell, Stephen A. Krawetz, Peter Snyder, Dana Ohl, Nanette Santoro, Esther Eisenberg, and Heping Zhang, for the NICHD Reproductive Medicine Network
Liu 2017	Chang Liu, Guimei Feng, Wei Huang, Qiuyi Wang, Shiyuan Yang, Jing Tan, Jing Fu and Dong Liu
Lord 2006	J Lord, R Thomas, B Fox, U Acharya and T Wilkin

Moll 2006	Etelka Moll, Patrick M MBossuyt, Johanna C Korevaar, Cornelis B Lambalk, and Fulco van der Veen,
Morin-Papunen 2012	Laure Morin-Papunen, Anni S. Rantala, Leila Unkila-Kallio, AilaTiitinen, MarittaHippeläinen, Antti Perheentupa, Helena Tinkanen, RistoBloigu, Katri Puukka, AimoRuokonen and Juha S. Tapanainen
Nazik 2012	Hakan Nazik and YakupKumtepe
Palomba 2005	Stefano Palomba, Francesco Orio, Jr., Angela Falbo, Francesco Manguso, Tiziana Russo, Teresa Cascella, Achille Tolino, Enrico Carmina, Annamaria Colao and Fulvio Zullo
Sahin 2004	Yılmaz Şahin, Ünal Yirmibeş, Fahrettin Keleştimur and Ercan Aygen
Vegetti 1999	W. Vegetti, A. Riccaboni, M. Columbo, E. Baroni, D. Diaferia, G. Ragni and P.G. Crosignani.
Williams 2009	C. D. Williams, L. M. Pastore, W. B.Shelly, A. P.Bailey, D. C.Baras and B. G.Bateman,
PCOSAct trial (Wu 2017)	Xiao-Ke Wu, ElisabetStener-Victorin, Hong-Ying Kuang, Hong-Li Ma, Jing-Shu Gao, Liang-Zhen Xie, Li-Hui Hou, Zhen-Xing Hu, Xiao-Guang Shao, Jun Ge, Jin-Feng Zhang, Hui-Ying Xue, Xiao-Feng Xu, Rui-Ning Liang, Hong-Xia Ma, Hong-Wei Yang, Wei-Li Li, Dong-Mei Huang, Yun Sun, Cui-Fang Hao, Shao-Min Du, Zheng-Wang Yang, Xin Wang, Ying Yan, Xiu-Hua Chen, Ping Fu, Cai-Fei Ding, Ya-Qin Gao, Zhong-Ming Zhou, Chi Chiu Wang, Tai-Xiang Wu, Jian-Ping Liu, Ernest H. Y. Ng, Richard S. Legro and Heping Zhang, for the PCOSAct Study Group

*Note: This publication included two studies, both of which were included.

Supplementary Table 5 Eunice Kennedy Shriver National Institutes of Child Health and Human Development, Reproductive Medicine Network

Names	**Affiliations**	**Support by NIH Grants**
Richard S. Legro, M.D.	Pennsylvania State University College of Medicine, Hershey, PA	U10 HD27049, U10 HD38992, U10 HD055925, U10 HD39005, U10 HD38998, U10 HD055936, U10 HD055942, U10 HD055944, U54 HD29834, UL1 TR000127, U01 HD38997, U10 HD27011, U10 HD33172, U10 HD38988, U10 HD38999, MO1RR00056, MO11RR10732, C06 RR016-499
Robert G. Brzyski, M.D., Ph.D.	University of Texas Health Science Center at San Antonio, San Antonio	
Michael P. Diamond, M.D.	Georgia Regents University, Augusta; Wayne State Univeristy, Detroit	
Christos Coutifaris, M.D., Ph.D.	University of Pennsylvania School of Medicine, Philadelphia	
William D. Schlaff, M.D.	University of Colorado, Aurora	
Peter Casson, M.D.	University of Vermont, Burlington	
Gregory M. Christman, M.D.	University of Michigan, Ann Arbor	
Hao Huang, M.D., M.P.H.	Yale University School of Public Health, New Haven, CT	
Qingshang Yan, Ph.D.	Yale University School of Public Health, New Haven, CT	
Ruben Alvero, M.D.	University of Colorado, Aurora	
Daniel J. Haisenleder, Ph.D.	Ligand Core Lab, Univ of Virginia Center for Research in Reproduction, Charlottesville	
Kurt T. Barnhart, M.D.	University of Pennsylvania School of Medicine, Philadelphia	
G. Wright Bates, M.D.	University of Alabama at Birmingham, Birmingham	
Rebecca Usadi, M.D.	Carolinas Medical Center, Charlotte, NC	
Scott Lucidi, M.D.	Virginia Commonwealth University, Richmond	
Valerie Baker, M.D.	Stanford University Medical Center, Stanford, CA	
J.C. Trussell, M.D.	State University of New York Upstate Medical University, Onondaga	

Stephen A. Krawetz, Ph.D.	Wayne State Univeristy, Detroit
Peter Snyder, M.D.	University of Pennsylvania School of Medicine, Philadelphia
Dana Ohl, M.D.	University of Michigan, Ann Arbor
Nanette Santoro, M.D.	University of Colorado, Aurora
Huiman X. Barnhart, Ph.D.	Duke University Medical Center, Durham, NC
Bruce R. Carr, M.D.	University of Texas Southwestern Medical Center, Dallas
Michael P. Diamond, M.D.	Wayne State Univeristy, Detroit
Sandra A. Carson, M.D.	Baylor College of Medicine, Houston
Michael P. Steinkampf, M.D.	University of Alabama, Birmingham
Peter G. McGovern, M.D.	University of Medicine and Dentistry of New Jersey, Newark
Nicholas A. Cataldo, M.D.	Stanford University, Stanford, CA
Gabriella G. Gosman, M.D.	University of Pittsburgh, Pittsburgh
John E. Nestler, M.D.	Virginia Commonwealth University School of Medicine, Richmond
Linda C. Giudice, M.D., Ph.D.	University of California at San Francisco, San Francisco
Phyllis C. Leppert, M.D., Ph.D.	National Institute of Child Health and Human Development, Bethesda, MD
Evan R. Myers, M.D., M.P.H.	Duke University Medical Center, Durham
Esther Eisenberg, M.D., M.P.H.	Fertility and Infertility Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, Rockville, MD
Heping Zhang, Ph.D.	Yale University School of Public Health, New Haven, CT