Management of Infertility
Management of Infertility

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
5600 Fishers Lane
Rockville, MD 20857
www.ahrq.gov

Contract No. 290-2015-00004-I

Prepared by:
Duke Evidence-based Practice Center
Durham, NC

Investigators:
Evan R. Myers, M.D., M.P.H.
Jennifer L. Eaton, M.D., M.S.C.I.
Kara A. McElligott, M.D., M.P.H.
Patricia G. Moorman, Ph.D., M.S.P.H.
Ranee Chatterjee, M.D., M.P.H.
Arthurine K. Zakama, M.D.
Karen Goldstein, M.D., M.S.P.H.
Jennifer Strauss, Ph.D.
Remy R. Coeytaux, M.D., Ph.D.
Adam Goode, D.P.T., Ph.D.
Ethan Borre, B.A.
Geeta K. Swamy, M.D.
Amanda J. McBroom, Ph.D.
Kathryn Lallinger, M.S.L.S.
Robyn Schmidt, B.A.
J. Kelly Davis, B.A.
Victor Hasselblad, Ph.D.
Gillian D. Sanders, Ph.D.

AHRQ Publication No. 19-EHC014-EF
May 2019
Key Messages

Purpose of Review
Evaluate the comparative effectiveness and safety of treatments for common causes of infertility.

Key Messages

- The ability to compare the effectiveness of treatments would be enhanced by greater consistency in reporting of outcomes, particularly live birth rates, as well as reporting of diagnosis-specific outcomes for treatments, such as assisted reproductive technology, that are used for multiple diagnoses.
- Letrozole most likely results in more live births with lower multiple births than clomiphene alone in women with polycystic ovary syndrome.
- For women with unexplained infertility, there is most likely shorter time to pregnancy for women with immediate in vitro fertilization (IVF) than for those who undergo other treatments prior to IVF. For the outcomes of live birth, multiple births, ectopic pregnancy, miscarriage, low birthweight, and ovarian hyperstimulation syndrome however, there may be no difference between the two groups.
- Across all diagnoses, elective single-embryo transfer results in slightly lower live birth rates but substantially lower reductions in multiple birth rates than multiple-embryo transfer.
This report is based on research conducted by the Duke Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-2015-00004-I. The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.

The information in this report is intended to help healthcare decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of healthcare services. This report is not intended to be a substitute for the application of clinical judgment. Anyone who makes decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report is made available to the public under the terms of a licensing agreement between the author and the Agency for Healthcare Research and Quality. This report may be used and reprinted without permission except those copyrighted materials that are clearly noted in the report. Further reproduction of those copyrighted materials is prohibited without the express permission of copyright holders.

AHRQ or U.S. Department of Health and Human Services endorsement of any derivative products that may be developed from this report, such as clinical practice guidelines, other quality enhancement tools, or reimbursement or coverage policies, may not be stated or implied.

This report may periodically be assessed for the currency of conclusions. If an assessment is done, the resulting surveillance report describing the methodology and findings will be found on the Effective Health Care Program website at www.effectivehealthcare.ahrq.gov. Search on the title of the report.

Persons using assistive technology may not be able to fully access information in this report. For assistance contact epc@ahrq.hhs.gov.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of systematic reviews to assist public- and private-sector organizations in their efforts to improve the quality of healthcare in the United States. These reviews provide comprehensive, science-based information on common, costly medical conditions, and new healthcare technologies and strategies.

Systematic reviews are the building blocks underlying evidence-based practice; they focus attention on the strength and limits of evidence from research studies about the effectiveness and safety of a clinical intervention. In the context of developing recommendations for practice, systematic reviews can help clarify whether assertions about the value of the intervention are based on strong evidence from clinical studies. For more information about AHRQ EPC systematic reviews, see www.effectivehealthcare.ahrq.gov/reference/purpose.cfm.

AHRQ expects that these systematic reviews will be helpful to health plans, providers, purchasers, government programs, and the healthcare system as a whole. Transparency and stakeholder input are essential to the Effective Health Care Program. Please visit the website (www.effectivehealthcare.ahrq.gov) to see draft research questions and reports or to join an email list to learn about new program products and opportunities for input.

If you have comments on this systematic review, they may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857, or by email to epc@ahrq.hhs.gov.

Gopal Khanna, M.B.A.  
Director  
Agency for Healthcare Research and Quality

Arlene S. Bierman, M.D., M.S.  
Director  
Center for Evidence and Practice Improvement  
Agency for Healthcare Research and Quality

Stephanie Chang, M.D., M.P.H.  
Director  
Evidence-based Practice Center Program  
Center for Evidence and Practice Improvement  
Agency for Healthcare Research and Quality

Suchitra Iyer, Ph.D.  
Task Order Officer  
Center for Practice Improvement  
Agency for Healthcare Research and Quality
Acknowledgments
The authors thank Megan von Isenburg, M.S.L.S., for help with the literature search and retrieval; Samantha Bowen, Ph.D., for project leadership; Megan Chobot, M.S.L.S., for assistance with project coordination; and Liz Wing, M.A., and Rebecca Gray, Ph.D., for editorial assistance.

Key Informants
In designing the study questions, the EPC consulted several Key Informants who represent the end-users of research. The EPC sought the Key Informant input on the priority areas for research and synthesis. Key Informants are not involved in the analysis of the evidence or the writing of the report. Therefore, in the end, study questions, design, methodological approaches, and/or conclusions do not necessarily represent the views of individual Key Informants.

Key Informants must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any conflicts of interest.

The list of Key Informants who provided input to this report follows:

Sheree Boulet, Dr.P.H., M.P.H.  
Centers for Disease Control and Prevention  
Atlanta, GA

Barbara Collura, M.A.  
RESOLVE: The National Fertility Association  
McLean, VA

Esther Eisenberg, M.D.  
National Institute of Child Health and Human Development  
Bethesda, MD

Tina Groat, M.D., M.B.A., FACOG  
UnitedHealthcare  
Minneapolis, MN

Dmitry Kissin, M.D., M.P.H.  
Centers for Disease Control and Prevention  
Atlanta, GA

Dan I. Lebovic, M.D.  
Center for Reproductive Medicine and Advanced Reproductive Technologies  
Minneapolis, MN

Richard S. Legro, M.D.  
Pennsylvania State University  
State College, PA

Bradley J. Van Voorhis, M.D.  
Carver College of Medicine  
University of Iowa  
Iowa City, IA
Technical Expert Panel

In designing the study questions and methodology at the outset of this report, the EPC consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

Technical Experts must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

The list of Technical Experts who provided input to this report follows:

Sheree Boulet, Dr.P.H., M.P.H.*
Centers for Disease Control and Prevention
Atlanta, GA

Barbara Collura, M.A.
RESOLVE: The National Fertility Association
McLean, VA

Esther Eisenberg, M.D.
National Institute of Child Health and Human Development
National Institutes of Health
Bethesda, MD

Audrey Gassman, M.D.*
Center for Drug Evaluation and Research
U.S. Food and Drug Administration
Silver Spring, MD

David Grainger, M.D., M.P.H.*
Center for Reproductive Medicine
University of Kansas School of Medicine
Wichita, KS

Tina Groat, M.D., M.B.A., FACOG
UnitedHealthcare
Minneapolis, MN

Dan I. Lebovic, M.D.*
Center for Reproductive Medicine and Advanced Reproductive Technologies
Minneapolis, MN

Richard S. Legro, M.D.
Pennsylvania State University
State College, PA

Jay Sandlow, M.D.*
Medical College of Wisconsin
Milwaukee, WI

Bradley J. Van Voorhis, M.D.*
Carver College of Medicine
University of Iowa
Iowa City, IA

*Provided input on Draft Report.
Peer Reviewers

Prior to publication of the final evidence report, EPCs sought input from independent Peer Reviewers without financial conflicts of interest. However, the conclusions and synthesis of the scientific literature presented in this report do not necessarily represent the views of individual reviewers.

Peer Reviewers must disclose any financial conflicts of interest greater than $5,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals with potential nonfinancial conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

The list of Peer Reviewers follows:

Kurt T. Barnhart, M.D., M.S.C.E.                      Brooke V. Rossi, M.D.
University of Pennsylvania Medical Center           School of Medicine
Philadelphia, PA                                         Case Western Reserve University
                                                        Cleveland, OH

Clarisa Gracia, M.D., M.S.C.E.                      John M. Thorp, Jr., M.D.
Perelman School of Medicine                        University of North Carolina
University of Pennsylvania                        Chapel Hill, NC
Philadelphia, PA

Philadelphia, PA
Management of Infertility

Structured Abstract

Objective. Previous studies have demonstrated varying success for treatment of infertility. Much of this literature, however, does not focus on treatment of women with specific diagnoses. This systematic review evaluated the comparative effectiveness and safety of fertility treatment strategies for (a) women of reproductive age (18–44) who are infertile due to polycystic ovary syndrome (PCOS), endometriosis, unknown reasons, or tubal or peritoneal factors or (b) couples with male factor infertility, and evaluated short- and long-term health outcomes of gamete donors in infertility.

Data sources. We searched PubMed®, Embase®, and the Cochrane Database of Systematic Reviews for English-language studies published from January 1, 2007, to October 3, 2018, that reported live birth rates, pregnancy and neonatal outcomes, time to pregnancy, and short-term and long-term adverse outcomes for mothers and children born after infertility treatment. For male and female donors, we searched for studies reporting short- and long-term adverse effects and quality-of-life outcomes.

Review methods. Two investigators screened each abstract and full-text article for inclusion; abstracted data; and performed quality ratings, applicability ratings, and evidence grading. Where appropriate, random-effects models were used to compute summary estimates of effects.

Results. We identified a total of 151 studies/primary articles that met our inclusion criteria: 56 for PCOS, 7 for endometriosis, 50 for infertility secondary to unknown causes, 8 for tubal/peritoneal factor infertility, 23 for male factor infertility, and 5 for outcomes in male and female gamete donors. There were also 21 studies that adjusted for cause of infertility but whose findings were relevant across all infertility diagnoses. For women with infertility associated with PCOS, there was moderate strength of evidence (SOE) that letrozole results in higher live birth rates than clomiphene while reducing multiple births and with no difference in ectopic pregnancies (moderate SOE). No differences were seen in low birthweight or time to pregnancy (low SOE). There was moderate SOE that there is no difference between clomiphene and metformin as primary therapy. Comparing laparoscopic ovarian drilling with oral agents, live birth rates were not different (moderate SOE). For couples with unexplained infertility, there is no difference between the oral agents of letrozole and anastrozole for the outcome of ectopic pregnancy (low SOE), but evidence is insufficient for other outcomes of interest. There was also no difference between differing adjunct treatments used in combination with oral agents and intrauterine insemination (IUI) for the outcomes of live birth, miscarriage, and ovarian hyperstimulation syndrome (OHSS) (low SOE for all outcomes). Time to pregnancy was shorter with immediate in vitro fertilization (IVF) compared with strategies that started with clomiphene and IUI or gonadotropins and IUI, followed by IVF if necessary (moderate SOE). For couples with male factor infertility, live birth rate (moderate SOE) and miscarriage (low SOE) did not differ between intracytoplasmic sperm injection (ICSI) and intracytoplasmic morphological sperm injection. (The latter is not used in the United States.) For oocyte donors, studies suggested a lower incidence of OHSS with gonadotropin-releasing hormone (GnRH) agonist trigger than with human chorionic gonadotropin (hCG) trigger (low SOE). However, there was a lack of evidence on any long-term outcomes. Evidence concerning specific comparisons was
insufficient for couples with tubal factor or endometriosis infertility. Findings applicable across all indications for infertility for couples undergoing assisted reproductive technology (ART) included lower live birth rates for African-Americans compared with other racial/ethnic groups (low SOE); lower live birth rates but significant reductions in multiple birth rates with elective single-embryo transfer compared with multiple embryo transfer (low SOE); no increase in most maternal cancers after ART treatment after adjustment for infertility in general or specific causes (low SOE); and, for children born after ART, a possible increased risk of neurodevelopmental disorders after ICSI compared with IVF (low SOE).

Conclusions. Although there is evidence supporting some strategies for treatment of infertility, both for specific diagnoses and for couples with any diagnosis, consensus on which outcomes to collect and report, and which areas of uncertainty are most important to resolve, is needed in order to design future studies that will improve the ability of patients and clinicians to make optimal decisions.
# Contents

Evidence Summary ................................................................................................................. ES-1

Introduction ............................................................................................................................ 1
  Background ............................................................................................................................. 1
  Condition ................................................................................................................................. 1
  Treatment Strategies ........................................................................................................... 1
  Complexity of Decision Making for Treatments of Infertility ........................................... 3
  Limitations of the Evidence Base ......................................................................................... 3
Scope and Key Questions ........................................................................................................ 4
  Scope of the Review ............................................................................................................... 4
  Key Questions ....................................................................................................................... 4
Organization of This Report .................................................................................................... 7

Methods .......................................................................................................................................... 8
  Topic Refinement and Review Protocol .................................................................................. 8
  Literature Search Strategy ...................................................................................................... 8
    Search Strategy .................................................................................................................... 8
    Inclusion and Exclusion Criteria .......................................................................................... 9
    Study Selection .................................................................................................................... 12
    Data Extraction .................................................................................................................. 13
  Quality (Risk of Bias) Assessment of Individual Studies ...................................................... 13
  Data Synthesis ...................................................................................................................... 14
  Strength of the Body of Evidence ......................................................................................... 15
  Applicability .......................................................................................................................... 16
  Peer Review and Public Commentary .................................................................................... 16

Results .......................................................................................................................................... 17
  Results of Literature Searches ............................................................................................... 17
  Description of Included Studies ............................................................................................ 19
  Key Question 1. PCOS ............................................................................................................ 19
    Description of Studies Included for KQ 1 (PCOS) .............................................................. 19
    Key Points for PCOS .......................................................................................................... 20
    Detailed Synthesis by Treatment for PCOS ...................................................................... 21
    Subgroups of Interest for PCOS ......................................................................................... 46
    Strength of Evidence for PCOS .......................................................................................... 46
  Key Question 2. Endometriosis ............................................................................................ 52
    Description of Included Studies for KQ 2 (Endometriosis) ................................................. 53
    Key Points for Endometriosis ............................................................................................ 53
    Detailed Synthesis by Treatment for Endometriosis .......................................................... 53
    Subpopulations of Interest for Endometriosis ................................................................... 55
    Strength of Evidence for Endometriosis .......................................................................... 55
  Key Question 3. Unexplained Infertility ................................................................................. 57
    Description of Included Studies for KQ 3 (Unexplained Infertility) ................................... 57
    Key Points for Unexplained Infertility .............................................................................. 57
    Detailed Synthesis by Treatment for Unexplained Infertility ............................................ 58
    Strength of Evidence for Unexplained Infertility ............................................................. 74
Table 37. Report findings and major guidelines/recommendations—KQ 3, unexplained infertility ................................................................. 111
Table 38. Report findings and major guidelines/recommendations—KQ 4, tubal and peritoneal factor infertility ........................................................................................................................... 112
Table 39. Report findings and major guidelines/recommendations—KQ 5, male factor infertility ..................................................................................................................................................... 114
Table 40. Report findings and major guidelines/recommendations—KQ 6, donors ........................................ 115
Table 41. Report findings and major guidelines/recommendations—all KQs: long-term outcomes of treatments .................................................................................................................................... 117
Table 42. Potential issues with applicability of included studies ............................................................ 119

Figures
Figure A. Analytic framework.............................................................................................................. ES-5
Figure 1. Analytic framework ........................................................................................................... 6
Figure 2. Literature flow diagram ...................................................................................................... 18
Figure 3. Effect of ICSI versus IMSI on live birth ......................................................................... 86

Appendixes
Appendix A. Exact Search Strings
Appendix B. Data Abstraction Elements
Appendix C. List of Included Studies
Appendix D. List of Excluded Studies
Appendix E. Characteristics of Included Studies
Appendix F. AMSTAR Quality Assessment for Systematic Reviews
Appendix G. Risk of Bias Assessment for Included Studies
Appendix H. Supplemental Project To Assess the Transparency of Reporting for Trials
Evaluating Treatment for Infertility
Evidence Summary

Background

Condition and Therapeutic Strategies

“Infertility” has traditionally been defined as failure to achieve pregnancy after 12 months of regular unprotected intercourse with the same partner (or after 6 months for women greater than 35 years of age). However, as many as half of such couples will conceive without intervention over the next 12-24 months. Because of this, the term “subfertility” is preferred by many.\(^1\) From a population perspective, couples who meet the dichotomous criteria for “infertility” include couples who are “normal” but who are in the upper end of the population distribution for “time to pregnancy,” and couples who have a physiological or anatomical cause for a prolonged time to pregnancy. However, to be concise, we will use the term “infertility” throughout this report.

Self-reported infertility in the United States, using the 12-month definition, affected approximately 6 percent of married women aged 15-44 in the 2006-2010 National Survey of Family Growth (the most recent available data).\(^2\) In one population-based study, approximately 10 percent of pregnant women reported receiving infertility treatment, with 29 percent of these women using fertility-enhancing medications; 21 percent using assisted reproductive technology (ART), including \textit{in vitro} fertilization (IVF); 15 percent using artificial insemination with fertility-enhancing drugs; and 23 percent using other treatments, including surgery.\(^3\) Other estimates of the prevalence of infertility treatment are similar.\(^4\)-\(^8\) Particularly in the United States, where availability of infertility services is variable depending on a number of factors, particularly insurance coverage, utilization of infertility treatments may underestimate the overall burden of infertility.

The most common demographic factor associated with female infertility is “advanced reproductive age,” although the probability of pregnancy begins to decline by the mid-20’s, the slope of decline sharply increases by age 35.\(^9\) Other common causes of female infertility include polycystic ovary syndrome (PCOS), endometriosis, occlusion of the fallopian tubes from prior infectious disease,\(^6\) and infertility secondary to cancer treatment.\(^10\)-\(^12\) Isolated male factor infertility affects approximately 17 percent of couples seeking treatment, with 34.6 percent of couples having both male and female diagnoses.\(^13\)

Treatment options are usually dependent on the underlying etiology of infertility. For female causes, options include surgical management of tubal occlusion, surgical treatment of endometriosis, ovarian “drilling” for treatment of PCOS, use of ovulation-induction agents including oral (clomiphene citrate or letrozole) and injected drugs (gonadotropins), artificial insemination with either partner or donor sperm (depending on partner fertility status), and ART, which includes both traditional IVF (fertilization of the egg by the sperm occurs without direct manipulation) and IVF with \textit{intra-cytoplasmic sperm injection} (ICSI) (fertilization occurs via direct injection of sperm into the egg).\(^14,15\) Treatment options for male factor infertility include medical treatment of a diagnosed endocrinopathy or other conditions affecting sperm production, empiric treatments with hormonal or other agents, surgical management of varicocele, intrauterine insemination, IVF, ICSI, or use of donor sperm.\(^16\) Options appropriate for some diagnoses (e.g., ovulation induction in PCOS or unexplained infertility) may not be appropriate for others (e.g., women with documented tubal occlusion). In other cases, the appropriate
comparisons may involve sequencing or combinations of treatment options—for example, one strategy might consist of several cycles of ovulation induction, followed by ART only if pregnancy does not occur, compared to proceeding directly to ART. Note that throughout this report, we use the term “adjunct treatments” to refer to interventions performed within a major treatment category (for example, comparison of metformin to placebo as pretreatment in women with PCOS undergoing IVF).

Although there has been ongoing debate about the most appropriate outcome for evaluation of infertility treatments, there is a growing consensus that live birth is the most important patient-centered outcome.\textsuperscript{17,18} Trade-offs between outcomes (particularly multiple gestations), time to pregnancy, and out-of-pocket costs might be different between the various treatment strategies even if cumulative live birth rates were identical.

Different treatments also carry different safety risks. There are known short-term risks such as ovarian hyperstimulation syndrome (OHSS) or acute risks associated with any surgery. Surgery may have additional longer-term risks which may affect subsequent fertility (such as scarring or decreased ovarian reserve with procedures such as laparoscopic ovarian drilling (LOD). The literature suggests that observed associations between infertility treatment and female reproductive cancers, particularly ovarian cancer, are likely the result of the underlying infertility rather than treatment itself. There is, however, some uncertainty surrounding some cancer outcomes in subgroups of patients.\textsuperscript{19-21}

Some adverse pregnancy outcomes, such as preterm birth, are associated with infertility treatment; however, many of the conditions associated with infertility are also associated with these adverse outcomes, complicating assessment of comparative effectiveness.\textsuperscript{22-25} There may also be direct effects of some treatments that have unclear implications for long-term health in children born after these treatments.\textsuperscript{26,27} Finally, infertility clearly has an emotional impact,\textsuperscript{12,28,29} and the comparative effects of infertility treatments on quality of life are an important consideration for both women and men.

There may be significant variation in outcomes of different treatments in specific subpopulations. For example, age affects the likelihood of conception, and the risk of many pregnancy complications associated with infertility treatments, such as preterm birth or low birthweight, are also increased with higher maternal age. Obesity is common in women with PCOS, and, like older maternal age, is also associated with adverse pregnancy outcomes independent of its association with infertility. The utilization and outcomes of infertility treatment differ among different racial and ethnic groups, even after adjusting for insurance coverage.\textsuperscript{30-33}

Finally, a unique subpopulation is women who donate oocytes for use by other couples in ART. There are almost no data on the long-term safety of multiple courses of ovulation induction for the purposes of oocyte donation.\textsuperscript{34} In addition, there are complex ethical and legal considerations, including the balance between fair compensation and inducement,\textsuperscript{35} and sharing information about donors with recipients.\textsuperscript{36}

**Scope and Key Questions**

This systematic review evaluates the comparative safety and effectiveness of available treatment strategies for women of reproductive age (18–44) who are infertile due to PCOS, endometriosis, unknown reasons, or tubal or peritoneal factors; the comparative safety and effectiveness of available treatment strategies for couples with male factor infertility; and the short- and long-term health outcomes of donors in infertility.
The specific Key Questions (KQs) addressed in this review are listed below, and Figure A displays the analytic framework that guided our work.

- **KQ 1.** What are the comparative safety and effectiveness of available treatment strategies for women with *polycystic ovary syndrome* who are infertile and who wish to become pregnant?
  - **KQ 1a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, body mass index (BMI), presence of other potential causes of female infertility, or presence of male factor infertility?
- **KQ 2.** What are the comparative safety and effectiveness of available treatment strategies for women with *endometriosis* who are infertile and who wish to become pregnant?
  - **KQ 2a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, stage of endometriosis, presence of other potential causes of female infertility, or presence of male factor infertility?
- **KQ 3.** What are the comparative safety and effectiveness of available treatment strategies for women who are infertile for *unknown reasons* and who wish to become pregnant?
  - **KQ 3a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, presence of other potential causes of female infertility, or presence of male factor infertility?
- **KQ 4.** What are the comparative safety and effectiveness of available treatment strategies for women with *tubal or peritoneal factors* *(e.g., pelvic adhesions)* who are infertile and who wish to become pregnant?
  - **KQ 4a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, presence of other potential causes of female infertility, or presence of male factor infertility?
- **KQ 5.** What are the comparative safety and effectiveness of available treatment strategies for couples with *male factor infertility* and no evidence of an underlying diagnosis associated with infertility in the female partner?
  - **KQ 5a.** Does the optimal treatment strategy vary by characteristics in either partner such as age, ovarian reserve, race, or BMI?
- **KQ 6.** What are the short- and long-term health outcomes of *donors in infertility*?
o KQ 6a. For female oocyte donors:
   1. Do specific aspects of the pre-donation evaluation identify potential donors at greater risk for short- or long-term adverse outcomes (e.g., OHSS, quality-of-life issues)?
   2. Do short- and long-term outcomes differ among different stimulation/retrieval protocols?

o KQ 6b. For male semen donors:
   - Are there long-term health, quality-of-life, or other adverse outcomes associated with donation?
Figure A. Analytic framework

Abbreviations: ART=assisted reproductive technology; BMI=body mass index; GnRH=gonadotropin-releasing hormone; KQ=Key Question; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome
Methods

Detailed methods are available in the full report and the posted protocol (http://effectivehealthcare.ahrq.gov/index.cfm).

Literature Search Strategy

To identify relevant published literature, we searched PubMed®, Embase®, and the Cochrane Database of Systematic Reviews, limiting the searches to studies conducted in adults and published from January 1, 2007, to October 3, 2018. Selection of the 2007 start date was based on establishing a one-year overlap with the search dates from a previous Agency for Healthcare Research and Quality (AHRQ) evidence report that assessed ART37 and input from Key Informants, who felt that the previous AHRQ review and more recent existing Cochrane reviews in this topic area would identify relevant high-quality studies. An experienced search librarian guided all searches. The exact search strings used are given in Appendix A of the full report.

We supplemented the electronic searches with a manual search of citations from a set of key primary and review articles. The reference lists for identified pivotal articles were manually hand-searched and cross-referenced against our database, and additional relevant articles not already under consideration were retrieved for screening. All citations were imported into an electronic bibliographical database (EndNote® Version X7; Thomson Reuters, Philadelphia, PA).

As a mechanism to ascertain publication bias in recent studies, we searched ClinicalTrials.gov to identify completed but unpublished studies (we also explored the possibility of publication bias in any quantitative synthesis of the included literature through meta-analysis techniques).

Approaches to identifying relevant gray literature included notification through the Federal Register to stakeholders, such as drug and device manufacturers, of the opportunity to submit scientific information packets. We also searched the ClinicalTrials.gov study registry and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) search portal to identify potentially relevant study records, and subsequently searched for relevant articles from completed studies.

We specified our inclusion and exclusion criteria based on the PICOTS (populations, interventions, comparators, outcomes, timing, and settings) identified for each question. For citations retrieved from PubMed, Embase, and the Cochrane Database of Systematic Reviews, two reviewers independently screened each title and abstract for potential relevance to the research questions using prespecified inclusion/exclusion criteria. Articles included by either reviewer underwent full-text screening. Articles meeting eligibility criteria at the full-text stage were included for data abstraction. Based on their clinical and methodological expertise, a pair of researchers were assigned to abstract data from each of the eligible articles. One researcher abstracted the data, and the second over-read the article and the accompanying abstraction to check for accuracy and completeness. Disagreements were resolved by consensus or by obtaining a third reviewer’s opinion if consensus could not be reached.

Risk of Bias Assessment of Individual Studies

We assessed methodological quality, or risk of bias, for each individual study using a components approach, assessing each study for specific aspects of design or conduct (such as allocation concealment for randomized controlled trials (RCTs), or use of methods to address
potential confounding), as detailed in AHRQ’s *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Briefly, we rated each study as being of good, fair, or poor quality based on its adherence to well-accepted standard methodologies. For each study, one investigator assigned a summary quality rating, which was then reviewed by a second investigator; disagreements were resolved by consensus or by a third investigator if agreement could not be reached.

We also rated quality for identified systematic reviews to provide additional context for the findings of the included studies. Rating was performed using AMSTAR (A Measurement Tool to Assess the Methodological Quality of Systematic Reviews). For each study, one investigator assigned a summary quality rating, which was then reviewed by a second investigator; disagreements were resolved by consensus or by a third investigator if agreement could not be reached. Reviews were then assigned overall quality scores of good (low risk of bias), fair (moderate risk of bias), or poor (high risk of bias). The consistency of the findings from these systematic reviews were incorporated into our strength of evidence ratings as described below.

**Data Synthesis**

We began by summarizing key features of the included studies for each KQ. To the degree that data are available, we abstracted information on study design; patient characteristics; clinical settings; interventions; and intermediate, final, and adverse event outcomes. If not reported, 95-percent confidence intervals for dichotomous outcomes (e.g., live birth rates) were calculated from the numbers provided in the study.

We then determined the feasibility of completing a quantitative synthesis (i.e., meta-analysis, decision analysis, or simulation model). For a meta-analysis, feasibility depends on the volume of relevant literature (requiring at least three relevant studies), conceptual homogeneity of the studies (similar intervention comparisons and outcome definitions), completeness of the reporting of results, and the adequacy and completeness of any existing meta-analyses.

**Strength of the Body of Evidence**

We graded the strength of evidence (SOE) for each outcome assessed using the approach described in the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. We also explored the consistency of our findings with recent systematic reviews and discussed agreement or disagreement, along with possible causes for disagreement and impact on strength of evidence ratings, in the results. A summary rating of high, moderate, or low strength of evidence was assigned for each outcome after discussion by two reviewers. When no evidence was available, or when evidence on the outcome was too weak, sparse, or inconsistent to permit any conclusion to be drawn, a grade of “insufficient” was assigned. This four-level rating scale consisted of the following definitions:

- **High Strength of Evidence**—We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable; i.e., another study would not change the conclusions.
- **Moderate Strength of Evidence**—We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- **Low Strength of Evidence**—We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous
deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.

- Insufficient Strength of Evidence—We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome.

Results

We briefly summarize the results of our literature searches, description of included studies, key points, and strength of evidence for each KQ. Note we only list here comparisons and outcomes with strength of evidence rated as low, moderate, or high. Full findings are available in the full report.

Summary of Studies

The literature search yielded 17,391 citations. In total, 1,909 studies were screened in full text, in which 1,748 were excluded for reasons listed in Figure 2 and Appendix D in the full report. We identified 161 articles describing 151 unique studies. The relationship of studies to the review questions is as follows: 56 studies relevant to KQ 1, 7 studies relevant to KQ 2, 50 studies relevant to KQ 3, 8 studies relevant to KQ 4, 23 studies relevant to KQ 5, and 5 studies relevant to KQ 6 (some studies were relevant to more than one KQ). There were also 21 studies relevant to findings across all KQs.

Key Question 1: PCOS

We identified 61 articles42-102 describing 56 studies that addressed the comparative safety and effectiveness of available treatment strategies for infertility in women with PCOS.

Key findings for outcomes in couples where the primary cause of infertility is PCOS include:

- Letrozole has a higher live birth rate than clomiphene citrate alone and lower multiple births, with no difference in ectopic pregnancy, or miscarriage (moderate for all outcomes), low birthweight, or time to pregnancy (low SOE for both these outcomes).
- Clomiphene citrate does not result in higher live birth rates compared with metformin (moderate SOE). Differences are also not found in the rates of multiple birth, ectopic pregnancy, or time to pregnancy (low SOE for all outcomes). There is a higher rate of miscarriage with combination clomiphene and metformin than clomiphene alone (low SOE)
- Letrozole or letrozole and berberine have a higher live birth rate than berberine alone (low SOE) with no difference in multiple births, miscarriage, or low birthweight rates (low SOE)
- There was no difference between clomiphene and tamoxifen for the outcomes of live birth or miscarriage (low SOE)
- There was no difference between laparoscopic ovarian drilling (LOD) and oral agents for live birth (moderate SOE) or miscarriage rates (low SOE). Multiple births were reduced given LOD (moderate SOE).
- Live birth (low SOE) and miscarriage rates (moderate SOE) did not differ between IVF treatment strategies.
- There was no difference in live birth rates for women who underwent lifestyle modification in combination with IVF compared with IVF alone (moderate SOE)
There was no difference between type 1 diabetes mellitus diagnoses in children conceived with ART compared to children conceived with no fertility treatment (moderate SOE).

As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE).

In general, SOE was judged insufficient or low for most outcomes, with the few exceptions including live births with the use of letrozole versus clomiphene or oral agents versus surgical management, and miscarriage between clomiphene and metformin or oral agents and surgical management which were rated moderate SOE. A common limitation across all comparisons was lack of precision for estimates of rare but important harms such as OHSS or surgical complications (Table A).

<p>| Table A. Summary of strength of evidence for major outcomes—KQ 1 (PCOS) |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| <strong>Comparison</strong>               | <strong>Outcome</strong>                 | <strong>Study Design</strong>            | <strong>Conclusion</strong>              | <strong>Strength of Evidence</strong>    |
|                             |                             | <strong>(Sample Size)</strong>           |                             | (Rationale)                 |
| Oral agents alone: Letrozole vs. | Live birth (any/patient)     | 1 RCT&lt;sup&gt;81&lt;/sup&gt; (644)    | <strong>Improvement</strong>: Letrozole or letrozole and berberine increase live birth rates compared to berberine alone. | Low (Imprecise, 1 study)    |
| Berberine vs. Berberine + Letrozole | Pregnancy complications: Multiple births | 1 RCT&lt;sup&gt;81&lt;/sup&gt; (644) | <strong>No difference</strong>: No significant difference between letrozole, berberine, or combination therapy | Low (Imprecise, 1 study)    |
| Oral agents alone: Letrozole vs. Clomiphene | Pregnancy complications: Miscarriage | 1 RCT&lt;sup&gt;81&lt;/sup&gt; (644) | <strong>No difference</strong>: No significant difference between letrozole, berberine, or combination therapy | Low (Imprecise, 1 study)    |
|                             | Neonatal outcomes: Birthweight | 1 RCT&lt;sup&gt;81&lt;/sup&gt; (644) | <strong>No difference</strong>: No significant difference between letrozole, berberine, or combination therapy | Low (Imprecise, one study)  |
| Oral agents alone: Letrozole vs. Clomiphene | Live birth (any/patient)     | 2 RCTs&lt;sup&gt;44,85&lt;/sup&gt; (909) | <strong>Improvement</strong>: Letrozole has higher live birth rates than clomiphene. | Moderate (Imprecise)        |
|                             |                             | 1 SR (9 studies, 1783 patients)&lt;sup&gt;103&lt;/sup&gt; |                             |                             |
| Pregnancy complications: Multiple births | 3 RCTs&lt;sup&gt;44,76,85&lt;/sup&gt; (886) | 1 SR (11 studies, 2385 patients)&lt;sup&gt;103&lt;/sup&gt; | <strong>Improvement</strong>: Letrozole has lower rates of multiple birth compared to clomiphene | Moderate (Inconsistent)    |
| Pregnancy complications: Ectopic pregnancy | 3 RCTs&lt;sup&gt;44,76,85&lt;/sup&gt; (886) | 1 SR (12 studies, 2385 patients)&lt;sup&gt;103&lt;/sup&gt; | <strong>No difference</strong>: No difference between letrozole and clomiphene. | Moderate (Imprecise)        |
| Pregnancy complications: Miscarriage | 3 RCTs&lt;sup&gt;44,76,85&lt;/sup&gt; (886) | 1 SR (12 studies, 2385 patients)&lt;sup&gt;103&lt;/sup&gt; | <strong>No difference</strong>: No statistical difference between letrozole and clomiphene | Moderate (Imprecise)        |
| Neonatal outcomes: Birthweight | 1 RCT&lt;sup&gt;44&lt;/sup&gt; (750) | <strong>No difference</strong>: No significant difference in birthweight between letrozole and clomiphene | Low (1 study)                |</p>
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral agents alone: Clomiphene vs. Metformin vs. Metformin + Clomiphene</td>
<td>Time to pregnancy</td>
<td>1 RCT(^44) (750)</td>
<td>No difference: No significant difference in time to pregnancy between clomiphene vs. letrozole</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Live birth (any/patient)</td>
<td>3 RCTs(^53,72,79) (842)</td>
<td>No difference: No statistical difference between clomiphene and metformin or between clomiphene and combination therapy of metformin and clomiphene</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>3 RCTs(^63,76,72) (921)</td>
<td>No difference: No differences in multiple birth rates between clomiphene alone, metformin alone, and clomiphene plus metformin</td>
<td>Low (Imprecise, suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>3 RCTs(^63,72,79) (1,005)</td>
<td>No difference: No difference between studied oral agents. Very few ectopic pregnancies overall.</td>
<td>Low (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs(^63,76,72,79) (817)</td>
<td>Increase: Higher rate of miscarriage in the combined therapy group (clomiphene and metformin) compared to clomiphene alone</td>
<td>Low (Suspected reporting bias, imprecise)</td>
</tr>
<tr>
<td></td>
<td>Time to Pregnancy</td>
<td>1 RCT(^53) (343)</td>
<td>No difference: No significant difference in time to pregnancy between clomiphene vs. metformin</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td>Oral agents alone: Clomiphene vs. Tamoxifen</td>
<td>Live birth (any/patient)</td>
<td>1 RCT(^99) (88)</td>
<td>No difference: No significant difference in live birth rates between tamoxifen and clomiphene</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT(^99) (88)</td>
<td>No difference: No significant difference in miscarriage rates between tamoxifen and clomiphene</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Active Acupuncture + Clomiphene vs. Control Acupuncture + Clomiphene vs. Active Acupuncture + Placebo vs. Control Acupuncture + Placebo</td>
<td>Live birth</td>
<td>1 RCT(^96) (1000)</td>
<td>Improvement: Live birth rates significantly higher for clomiphene vs. placebo; not significantly different for active vs. control Acupuncture</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT(^96) (1000)</td>
<td>No difference: no significant difference in ectopic pregnancy rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT(^96) (1000)</td>
<td>No difference: no significant difference in miscarriage rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>----------------------------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital Abnormalities</td>
<td>1 RCT (1000)</td>
<td>No difference: no significant difference in congenital abnormality rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
<td></td>
</tr>
<tr>
<td>Neonatal Death</td>
<td>1 RCT (1000)</td>
<td>No difference: no significant difference in neonatal death rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
<td></td>
</tr>
<tr>
<td>Oral agents alone vs. LOD</td>
<td>Live birth (any/patient)</td>
<td>1 SR (8 studies, 1,034 women)</td>
<td>No difference: No statistically significant differences between LOD and oral agents.</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 SR (15 studies, 1,129 women)</td>
<td>Reduction: There was a reduction in multiple births given LOD as compared to oral agents.</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT (80)</td>
<td>No difference: No significant differences in miscarriage between LOD and oral agents.</td>
<td>Low (Imprecise, suspected reporting bias)</td>
</tr>
<tr>
<td>Clomiphene citrate vs. low-dose FSH</td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>3 RCTs (1072)</td>
<td>No difference: Ectopic pregnancy rate did not differ between FSH and clomiphene strategies.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Lifestyle modifications + IVF vs. IVF alone</td>
<td>Live birth</td>
<td>3 RCTs (1688)</td>
<td>No difference: No difference in live birth rates for women who underwent lifestyle modification in combination with IVF compared with IVF alone</td>
<td>Moderate (Heterogeneity in interventions)</td>
</tr>
<tr>
<td>ART IVF: GnRH agonist +/- IVF vs. GnRH antagonist +/- IVF</td>
<td>Live birth (cycle)</td>
<td>4 RCTs (408)</td>
<td>No difference: No significant difference in included studies but varying interventions and comparators with low numbers of live birth</td>
<td>Low (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs (279)</td>
<td>No difference: No differences in miscarriage rates for GnRH agonist vs. antagonist, or hCG medium, hCG-free medium with transfer, and hCG-free medium without transfer.</td>
<td>Moderate (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>ART IVF: Fresh vs. Frozen Embryos in IVF for PCOS</td>
<td>Live birth (any/cycle)</td>
<td>1 RCT (1508)</td>
<td>Improvement: Live birth rates were significantly higher with frozen embryo transfer compared to fresh embryos.</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT (1508)</td>
<td>No difference: No difference in multiple births with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT (1508)</td>
<td>Reduction: Ectopic pregnancies were reduced with frozen embryo transfer</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT (1508)</td>
<td>Reduction: Miscarriages were reduced with frozen embryo transfer</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td>Neonatal Outcomes: Congenital abnormalities</td>
<td>1 RCT (1508)</td>
<td>No difference: No difference congenital abnormalities with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
<td></td>
</tr>
<tr>
<td>Neonatal Death</td>
<td>1 RCT (1508)</td>
<td>No difference: No difference neonatal deaths with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
<td></td>
</tr>
</tbody>
</table>
Comparison Outcome Study Design (Sample Size) Conclusion Strength of Evidence (Rationale)a

ART vs. no infertility treatment Long-term outcomes: Child (type 1 diabetes mellitus) 1 Obs93 (565,116 pregnancies) No difference: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with PCOS infertility conceived with ART compared to children conceived with no fertility treatment Moderate (Imprecise)

ART: IVF/ICSI vs. no treatment Live birth 1 Obs111 (69,028 cycles) Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent 2 embryo transfer (51.5%) as compared with single embryo transfer (46.6%) (p<0.0001). Low (Imprecise)

IUI vs. ART Live birth 1 Obs52 (19,884) Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent ART than those who used IUI Low (1 study)

aCriteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Key Question 2. Endometriosis

We identified seven individual studies that addressed infertility treatment for women with endometriosis.91,92,108-112

Key findings for couples where the primary cause of infertility is endometriosis in the female partner included:

- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)
- The live birth rate per cycle was higher in couples who underwent ART than those who used intrauterine insemination (IUI) (low SOE)
- SOE was rated insufficient for all other comparisons/outcomes.

In general, the SOE across all outcomes was judged to be insufficient or low, primarily due to imprecision and small numbers of studies, especially for both short-term harms (such as OHSS) (Table B).

Table B. Summary of strength of evidence for major outcomes—KQ 2 (endometriosis)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)a</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART: IVF/ICSI vs. no treatment</td>
<td>Live birth</td>
<td>1 Obs111 (69,028 cycles)</td>
<td>Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent 2 embryo transfer (51.5%) as compared with single embryo transfer (46.6%) (p&lt;0.0001).</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>IUI vs. ART</td>
<td>Live birth</td>
<td>1 Obs52 (19,884)</td>
<td>Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent ART than those who used IUI</td>
<td>Low (1 study)</td>
</tr>
</tbody>
</table>

aCriteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: ART=assisted reproductive technology; ICSI=intracytoplasmic sperm injection; IUI=intrauterine insemination; IVF=in vitro fertilization; KQ=Key Question; Obs=observational study
Key Question 3. Unexplained Infertility

We identified 50 individual studies that met inclusion criteria for KQ 3 and had unexplained infertility (infertility with no other documented female or male diagnosis).75,91,92,111,113-158

Key findings for couples with unexplained infertility included:

- There is no difference between the oral agents of letrozole and anastrozole for the outcome of ectopic pregnancy (low SOE) but evidence is insufficient for other outcomes of interest.
- There is no difference between letrozole and clomiphene for outcomes of multiple births or miscarriage (moderate SOE).
- There is no difference between differing adjunct treatments used in combination with oral agents and IUI for the outcomes of live birth, miscarriage, and OHSS (low SOE for all outcomes).
- There are no differences between immediate IVF versus other treatments prior to IVF for the outcomes of live birth, multiple births, ectopic pregnancy, miscarriage, low birthweight, and OHSS (low SOE for all outcomes). There is however shorter time to pregnancy with immediate IVF (moderate SOE).
- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)

SOE for most outcomes was judged to be insufficient or low, primarily because of imprecision or small numbers of studies of fair quality. Two exceptions were multiple births and miscarriages for oral agents without IUI where an existing systematic review existed, and time to pregnancy between different strategies for sequencing treatment, where precision was reasonable. In both cases SOE for these outcomes was judged to be moderate (Table C).

Table C. Summary of strength of evidence for major outcomes—KQ 3 (unexplained infertility)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Agents Without IUI</td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>2 RCTs132,157 (1,168)</td>
<td>No difference: No difference between letrozole and anastrozole:</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 SR159 (5 studies, 395 patients)</td>
<td>No difference: No difference between letrozole and clomiphene citrate</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs113,132,157 (1,248)</td>
<td>No difference: No difference between letrozole and clomiphene citrate</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 SR159 (5 studies, 395 patients)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clomiphene Citrate vs. Expectant Management</td>
<td>Pregnancy complications: Ectopic Pregnancy</td>
<td>2 RCTs136,149 (781)</td>
<td>No difference: No significant difference in ectopic pregnancy rates between clomiphene and expectant management</td>
<td>Low (Imprecise, heterogeneous interventions)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>2 RCTs136,149 (781)</td>
<td>No difference: No significant difference in ectopic pregnancy rates between clomiphene and expectant management</td>
<td>Low (Imprecise, heterogeneous interventions)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>---------------------------</td>
<td>------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Oral Agents vs. Unstimulated IUI vs. Expectant Management</td>
<td>Live birth</td>
<td>1 SR(^{160}) (3 studies, 370)</td>
<td>Improvement: A significant increase in live births was found for women treated with IUI and ovarian hyperstimulation compared to women treated with IUI only</td>
<td>Low (Inconsistent)</td>
</tr>
<tr>
<td>Adjunct Treatments with Oral Agents and IUI</td>
<td>Live birth</td>
<td>5 RCTs(^{139,149,156,157,158}) (1859)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>Adjunct Treatments with Oral Agents and IUI</td>
<td>Pregnancy complications: Miscarriage</td>
<td>5 RCTs(^{139,149,144,156,157}) (1859)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>Adjunct Treatments with Oral Agents and IUI</td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>3 RCTs(^{124,138,156}) (1189)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>Oral Agents With IUI vs. Gonadotropins With IUI</td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs(^{144,152,155}) (1,654)</td>
<td>No difference: No difference between oral agents with IUI versus gonadotropins with IUI</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Oral Agents With IUI vs. Gonadotropins With IUI</td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT(^{144}) (742)</td>
<td>Increased risk: Greater multiple gestations with gonadotropins compared to either clomiphene or letrozole</td>
<td>Low (one study)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Live birth</td>
<td>3 RCTs(^{118,120,131,151}) (812)</td>
<td>No difference: Live birth does not differ between differing strategies of other treatments prior to IVF</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Pregnancy complications: Multiple births</td>
<td>2 RCTs(^{118,131}) (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>3 RCTs(^{118,120,131,151}) (812)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs(^{118,120,131,151}) (812)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Neonatal outcomes: Birthweight</td>
<td>2 RCTs(^{118,131}) (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Time to pregnancy</td>
<td>2 RCTs(^{118,131}) (657)</td>
<td>Reduction: Shorter time to pregnancy with immediate IVF compared with other treatments prior to IVF.</td>
<td>Moderate</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 RCTs(^{118,131}) (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>ART: IVF vs. ICSI</td>
<td>Neonatal outcomes: Birth weight</td>
<td>1 Obs(^{31}) (90,401 cycles)</td>
<td>No difference: No significant differences in rates of low birth weight between ICSI versus conventional-IVF cycles</td>
<td>Low (1 study with moderate study limitations)</td>
</tr>
<tr>
<td>ART: Unspecified</td>
<td>Long-term outcomes: Child (cancer)</td>
<td>1 Obs(^{135}) (33,840)</td>
<td>No difference: The overall cancer incidence was not elevated in children born after assisted conception for unexplained infertility.</td>
<td>Low (Moderate study limitations)</td>
</tr>
</tbody>
</table>
Key Question 4. Tubal and Peritoneal Factor Infertility

We identified eight individual studies\textsuperscript{90,91,111,161-165} that addressed outcomes after treatment for tubal or peritoneal factor infertility.

Key findings for patients with tubal or peritoneal factor infertility included:

- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)
- The live birth rate was lower in women undergoing ICSI as compared to conventional IVF (low SOE)
- There was no difference between type 1 diabetes mellitus diagnoses in children born to patients with tubal factor infertility conceived with ART compared to children conceived with no fertility treatment (moderate SOE)
- SOE was rated insufficient for all other comparisons/outcomes.

The SOE was judged to be insufficient for most outcomes primarily due to imprecision based on few studies meeting our inclusion criteria (Table D).

Table D. Summary of strength of evidence for major outcomes—KQ 4 (tubal and peritoneal factor infertility)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design and Sample Size</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART: 2-embryo transfer vs. 1-</td>
<td>Live birth (patient)</td>
<td>1 Obs\textsuperscript{111} (69,028 cycles)</td>
<td>Improvement. The live birth rate per cycle was higher in couples who underwent 2 embryo transfer as compared with single embryo transfer</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>embryo transfer</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ART: IVF+ICSI vs. IVF</td>
<td>Neonatal outcomes: Birth weight</td>
<td>1 Obs\textsuperscript{91} (90,401 cycles)</td>
<td>No difference: No significant differences in rates of low birth weight between ICSI versus conventional-IVF cycles</td>
<td>Low (1 study with moderate study limitations)</td>
</tr>
<tr>
<td>ART vs. no fertility treatment</td>
<td>Long-term outcomes: Child (type 1 diabetes mellitus)</td>
<td>1 Obs\textsuperscript{90} (565,116 pregnancies)</td>
<td>No difference: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with tubal factor infertility conceived with ART compared to children conceived with no fertility treatment</td>
<td>Moderate (Imprecise)</td>
</tr>
</tbody>
</table>

Abbreviations: ART=assisted reproductive technology; ICSI=intra-cytoplasmic sperm injection; IVF=\textit{in vitro} fertilization; KQ=Key Question; Obs=observational study

\textsuperscript{a}Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.
Key Question 5. Male Factor Infertility

We identified 23 individual studies\(^\text{75,90-92,111,130,147,153,166-179}\) that addressed the comparative effectiveness or safety of interventions applied to patients with male factor infertility.

Key findings for patients with male factor infertility included:

- Live birth rate (moderate SOE) and miscarriage (low SOE) did not differ between intracytoplasmic sperm injection (ICSI) and intracytoplasmic morphological sperm injection (IMSI). Of note, IMSI is not used in the United States.
- There was no difference in live birth rates or any adverse pregnancy events between couples using frozen embryo versus fresh embryo transfer (low SOE).
- The overall cancer incidence was not elevated in children born after assisted conception for male factor infertility (low SOE).
- There was no difference between type 1 diabetes mellitus diagnoses in children born to patients with male factor infertility conceived with ART compared to children conceived with no fertility treatment (moderate SOE).
- Live birth rate (low SOE) improved with vitamin E or zinc supplementation relative to placebo or no supplementation.
- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE).

The SOE was judged to be insufficient or low for all outcomes except for the comparison of IVF versus ICSI for live birth and long term outcomes related to diabetes (Table E).

Table E. Summary of strength of evidence for major outcomes—KQ 5 (male factor infertility)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design and Sample Size</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ART IVF: ICSI or assisted hatching (1 embryo transferred) vs. ICSI or assisted hatching (multiple embryos transferred)</td>
<td>Live birth</td>
<td>2 Obs(^\text{111,171}) (272,717 cycles)</td>
<td>Improvement: Greater live births with multiple embryos transferred compared to 1 embryo transferred</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>TESE vs. ejaculated OAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ART IVF: Frozen vs. fresh embryo transfer</td>
<td>Live birth</td>
<td>1 RCT(^\text{177}) (2,157 patients)</td>
<td>No difference: no difference in live birth rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT(^\text{177}) (2,157 patients)</td>
<td>No difference: no difference in ectopic pregnancy rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design and Sample Size</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)a</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>-----------------------------</td>
<td>------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT(^{177}) (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in multiple birth rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT(^{177}) (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in miscarriage rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT(^{177}) (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in low birthweight rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>1 RCT(^{177}) (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in congenital anomalies rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (1 study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>IVF vs. ICSI</td>
<td>Live birth</td>
<td>3 RCTs(^{166,170,173}) (497 patients) 2 Obs(^{168,172}) (771,661 cycles)</td>
<td><strong>No difference</strong>, Meta-analysis of 3 RCTs does not demonstrate a difference between ICSI and IMSI.</td>
<td>Moderate (Moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT(^{166}) (121 patients) 1 Obs(^{168}) (499,135 cycles) 1 SR(^{180}) (6 studies, 552 women)</td>
<td><strong>No difference</strong>, Both included studies and an existing systematic review supported no difference in miscarriage. SOE was reduced because of quality of included studies and imprecision of findings.</td>
<td>Low (High study limitations, imprecise)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT(^{166}) (121 patients) 3 Obs(^{91,168,172}) (862,062 cycles)</td>
<td><strong>No difference</strong>: No significant differences in rates of low birth weight between ICSI versus conventional-IVF cycles</td>
<td>Low (Moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>ART: Unspecified</td>
<td>Long-term outcomes: Child (cancer)</td>
<td>1 Obs(^{121}) (924,427 patients)</td>
<td><strong>No difference</strong>: The overall cancer incidence was not elevated in children born after assisted conception for male factor infertility.</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>ART: Unspecified</td>
<td>Long-term outcomes: Child (type 1 diabetes mellitus)</td>
<td>1 Obs(^{90}) (565,116 pregnancies)</td>
<td><strong>No difference</strong>: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with male factor infertility conceived with ART compared to children conceived with no fertility treatment</td>
<td>Moderate (Imprecise)</td>
</tr>
<tr>
<td>Other strategies: Antioxidant use for Male Infertility</td>
<td>Live birth</td>
<td>1 SR(^{181}) (4 studies of 277 couples)</td>
<td><strong>Improvement</strong>: Increase in live birth rate associated with vitamin E or zinc supplementation relative to placebo or no supplementation</td>
<td>Low (Imprecise, small studies)</td>
</tr>
</tbody>
</table>

*aCriteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.*
Key Question 6. Donors in Infertility

We identified one fair-quality RCT\textsuperscript{182} and four retrospective observational studies, three fair-quality,\textsuperscript{183-185} and one poor-quality,\textsuperscript{186} that addressed short- or long-term health outcomes of donors in infertility.

Key findings for outcomes of sperm and oocyte donors included:

- For oocyte donors, observational studies suggest a lower incidence of OHSS with GnRH agonist trigger than with human chorionic gonadotropin (hCG) trigger (low SOE). However, there was a lack of evidence on any long-term outcomes.

Table F summarizes the SOE for KQ 6 and specifically for the incidence of OHSS with GnRH agonist trigger versus hCG trigger. All other short- and long-term outcomes had insufficient SOE or were not evaluated in the limited set of included studies.

Table F. Summary of strength of evidence for major outcomes—KQ 6 (donors in infertility)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)\textsuperscript{a}</th>
</tr>
</thead>
<tbody>
<tr>
<td>GnRH agonist (leuprolide acetate) vs. hCG trigger</td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 Obs\textsuperscript{183,184} (3824)</td>
<td>Improvement: Lower incidence of OHSS with GnRH agonist trigger than with hCG trigger.</td>
<td>Low (Moderate study limitations, imprecise)</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Findings Applicable Across All Infertility Diagnoses

We identified 26 articles\textsuperscript{21,167,187-210} described in 21 studies that addressed outcomes after treatment for infertility and adjusted for cause of infertility and therefore were considered relevant across all infertility diagnoses.

Findings applicable across all KQs for patients who undergo IVF/ICSI include:

- Clomiphene or gonadotropins ever use was not associated with increased risk of maternal cancer (low SOE).
- Women who undergo IVF demonstrated an increased risk of ovarian neoplasms and colorectal malignancies (low SOE) compared to women who do not undergo IVF. There is no evidence of a difference in invasive ovarian cancers (low SOE).
- For children born after ART, ICSI may be associated with an increased risk of autism compared to IVF (low SOE).
- In the United States, live birth rates after IVF/ICSI are lower for African-Americans than for other racial/ethnic groups after adjusting for other prognostic factors (low SOE).
- Elective single-embryo transfer is associated with lower live birth rates but a significant reduction in multiple birth rates compared to multiple-embryo transfer (low SOE for both outcomes).
- There was no difference in the odds of low birth weight between ICSI versus conventional IVF cycles (low SOE). However, among couples undergoing ART with a singleton pregnancy, frozen embryo transfers result in a higher average birthweight, with a subsequent reduction in the incidence of low birthweight and an increase in the incidence of macrosomia (low SOE).

Table G summarizes the SOE for findings that are applicable across all infertility diagnoses.

**Table G. Summary of strength of evidence for major outcomes—all infertility diagnoses**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clomiphene citrate and gonadotropin</td>
<td>Long-term outcomes: Maternal cancer</td>
<td>1 Obs(^9) (9892 patients)</td>
<td><strong>No difference.</strong> Ever use of clomiphene citrate was not statistically significantly associated with maternal ovarian, breast, endometrial, lung, thyroid, colon, or melanoma cancer. Gonadotropin use was not associated with increased risk for breast or endometrial cancer.</td>
<td>Low (Size of cohort not sufficient to detect modest increases in risk)</td>
</tr>
<tr>
<td>ART: IVF</td>
<td>Live birth (by race)</td>
<td>1 Obs(^{331}) (13,473 cycles)</td>
<td><strong>Greater disparity.</strong> Lower live birth rate for blacks as compared to white (p&lt;0.001)</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td></td>
<td>Live birth (by number of embryos transferred)</td>
<td>1 Obs(^{111}) (69,028 cycles)</td>
<td><strong>Improvement.</strong> Increased live birth rate per cycle with 2 embryo transfer as compared to single embryo transfer</td>
<td>Low (Imprecise, findings with moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births (by number of embryos transferred)</td>
<td>1 Obs(^{111}) (69,028 cycles)</td>
<td><strong>Greater risk.</strong> Multiple live birth rates are significantly higher with a 2-embryo transfer than a single embryo transfer, but do not increase further with 3- or 4-embryo transfers</td>
<td>Low (Imprecise, findings with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 Obs(^{93}) (8,948)</td>
<td><strong>No difference:</strong> No significant difference in rates of low birthweight using ART by assisted hatching, source of oocytes/semen, number of embryos or ICSI</td>
<td>Low (Imprecise)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital Anomalies</td>
<td>1 Obs(^{97}) (64,861)</td>
<td><strong>Greater risk.</strong> Risk of birth defects was greater in infants conceived using ART</td>
<td>Low (1 study)</td>
<td></td>
</tr>
<tr>
<td>Long-term outcomes: Child (Autism)</td>
<td>1 Obs(^{98}) (42,383)</td>
<td><strong>Greater risk.</strong> Risk of autism was greater in children conceived with ART with ICSI as compared to ART without ICSI</td>
<td>Low (Imprecise)</td>
<td></td>
</tr>
</tbody>
</table>
Comparison | Outcome | Study Design (Sample Size) | Conclusion | Strength of Evidence (Rationale) 
--- | --- | --- | --- | ---
Long-term outcomes: Maternal (cancer) | 2 Obs\(^{87,269}\) (280,950) | Greater risk. IVF was associated with a statistically significant increased risk of all ovarian neoplasms and borderline ovarian tumors, and colorectal cancer | Low (Imprecise, older study) | 
No difference: IVF however was not associated with an increased risk of invasive ovarian cancer, or melanoma | 

IVF+ICSI vs. IVF | Neonatal outcomes: Birth weight | 1 Obs\(^{91}\) (90,401 cycles) | No difference: No significant difference in the odds of low birth weight between ICSI versus conventional-IVF cycles | Low (1 study with moderate study limitations) | 

---

Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: ART=assisted reproductive technology; ICSI=intracytoplasmic sperm injection; IUI=intrauterine insemination; IVF=intravenous fertilization; KQ=Key Question; Obs=observational study

**Discussion**

In this Comparative Effectiveness Review, we reviewed 151 studies described in 161 publications that directly compared infertility management strategies in couples with infertility due to PCOS (KQ 1) or endometriosis (KQ 2); unexplained infertility (KQ 3); tubal and peritoneal factor infertility (KQ 4); and male factor infertility (KQ 5). We also explored the comparative safety and effectiveness of management strategies for donors in infertility (KQ 6).

Although the ultimate goal with any infertility management strategy is to improve live birth rates of healthy infants to a healthy couple, many studies initially identified in our review only reported on pregnancy rates or focused on other short-term outcomes and did not differentiate by the underlying causes of infertility. Our findings are based on those 151 studies which evaluated the comparative effectiveness of infertility management strategies in couples with a known cause of infertility (including unexplained infertility) and which evaluated the outcome of live birth or another long-term outcome.

**Findings in Relation to What Is Already Known**

The 2008 AHRQ Evidence Report on “Effectiveness of ART”\(^{37}\) found that approximately 80 percent of the 478 included studies were performed outside the United States, and that the majority of RCTs did not report delivery rates and obstetric outcomes. In that review, most studies did not have sufficient power to detect clinically meaningful differences in live birth rates, and had still lower power to detect differences in less frequent outcomes such as multiple births and complications. In addition, the previous report focused on outcomes of specific treatments (ovulation induction, superovulation, and IVF/ICSI) rather than a wider range of potential treatments, and infertility diagnosis was considered as subgroup analyses, rather than the primary basis for comparing treatments.

Methods for evidence synthesis, in particular for rating strength of evidence, have also been revised since that report. Although an increasing number of studies are using live birth rate as the...
primary outcome, the majority of the literature, particularly randomized trials, is still based on pregnancy or ongoing pregnancy. Lack of precision for comparative estimates of rates for less common but important outcomes, such as complications, continues to be a major limitation.

We compared our findings to evidence-based guidelines from the National Institute for Health and Care Excellence (NICE) in the United Kingdom (UK), and the American Society for Reproductive Medicine (ASRM). In general, findings of our review were concordant with the guidelines, with differences primarily attributable to differences in inclusion/exclusion criteria.

For women with PCOS, both NICE and ASRM support use of clomiphene citrate alone as first-line therapy, with the NICE guidance recommending ultrasound monitoring for dose adjustment to minimize risk of multiple pregnancy, followed by combination therapy with metformin or gonadotropins for women who do not conceive after a 3-6 month course of clomiphene alone. Both our review and NICE suggest letrozole may be superior to clomiphene as first line therapy, and that pretreatment with metformin may improve outcomes in women with PCOS being treated with gonadotropins.

For women with endometriosis, ASRM concluded that evidence for surgical treatment of women with mild to moderate endometriosis was insufficient to recommend treatment, while the NICE guidance suggests some benefit, and our review was inconclusive. For those patients going directly to ART, surgical treatment of endometriosis, including endometrioma, prior to ART does not improve outcomes.

For women with unexplained infertility, NICE recommends against use of oral agents entirely, while ASRM suggests clomiphene plus IUI may improve cycle fecundity compared to expectant management; our review found insufficient evidence. Based on our review, immediate IVF results in higher live birth rates and shorter time to pregnancy in women aged 38-42 compared with a trial of clomiphene and IUI or gonadotropins and IUI, with most live births ultimately resulting from IVF.

For women with suspected tubal factor infertility, both NICE and ASRM recommend imaging for diagnosis (which is outside the scope of our review), although, when ART is readily available and affordable, proceeding directly to ART without a definitive diagnosis of tubal disease may be more efficient.

For male factor infertility, our review found no relevant findings compared to the recommendations, primarily because of limited data on live birth outcomes.

For both male and female donors, both NICE and ASRM recommend psychological evaluation and counseling, including, for females, the short term risks of ovarian stimulation and oocyte collection; our review found evidence on outcomes was limited only to the known short-term risks of these procedures, with no evidence on potential longer term risks.

For long-term outcomes in women and children after infertility treatment, our review found limited or inconsistent evidence. Risks of adverse longer term maternal cancer outcomes were generally not increased after adjustment for the risk associated with infertility itself. ICSI however may be associated with an increased risk of neurodevelopmental disorders in children compared to those conceived through IVF. The NICE guidance was generally consistent with this assessment, and recommended that patients should be informed that any absolute risk was low, while there was still uncertainty about longer-term outcomes.

In general, our current review’s findings are consistent with the NICE and ASRM guidelines—there is a general consensus that the overall body of evidence for many aspects of infertility treatment across all patient groups is limited. One consistent limitation is the relative paucity of studies utilizing live birth per couple as the primary outcome.
Applicability

Two broad issues relate to the overall applicability of the available evidence to clinical practice in the United States—one geographic and one temporal. Many of the RCTs meeting our criteria were performed outside of the United States. Leaving aside any issues related to differences in study oversight or reporting, the populations of these studies may differ from U.S. infertility patients in two potentially important ways.

The first issue is that there may be clinically relevant differences between populations in terms of non-clinical factors affecting outcomes. For example, live birth rates for African-American women undergoing ART in the US are lower than for white women, which may reflect issues related to socioeconomic status, insurance coverage, or other factors (such as well-established racial differences in the risk of many adverse pregnancy outcomes). Differences in access to infertility services between countries may lead to differences in the likelihood of treatment success. Although the estimate of any relative difference between two interventions derived from an unbiased RCT should in theory be independent of the probability of specific outcomes, the more clinically relevant absolute difference may be substantially different (e.g., the risk of preterm birth in African-American compared to white women is consistently elevated). To the extent that the probability of specific outcomes of interest may differ between populations because of differences in genetic risk, exposures to other factors affecting risk, or non-biological factors such as access to care, there may be substantial differences in estimates of absolute risk differences. For relatively uncommon but important outcomes, these differences might also affect precision of estimates—confidence intervals for any treatment effect will be wider in populations where the outcome is less common.

In addition to the potential impact of race/ethnicity, there may be important differences in the distribution of socioeconomic status between populations. Access to infertility diagnosis and treatment varies across countries, and certainly within the United States. Differences in socioeconomic status could affect applicability in several ways. Differences in access to care may lead to differences in the spectrum of severity of “disease” for U.S. patients who given the financial burden of treatment options they may wait longer to undergo evaluations. Although summary statistics of baseline characteristics may allow some judgment of comparability, there may be potentially important differences in the distribution that are obscured by the typical reporting of means and standard deviations (particularly if the underlying characteristic is not normally distributed), or by differences within a given stage. Socioeconomic status may also potentially affect some important outcomes independently of any specific treatment—for example, neurodevelopmental outcomes such as specific learning skills may be strongly correlated with parental socioeconomic status.

The second issue is that changes in practice over time have a major impact on applicability, particularly for long-term outcomes. The long lag time between exposure to infertility treatment and the potential development of longer term outcomes such as cancer means that data available today necessarily reflect women exposed to treatments at least 10 years in the past; even if the specific exposure is similar, there may be differences between past and current practice in potentially important attributes such as dosage, timing, patient selection criteria, use of adjunctive treatments, etc. For example, evidence that immediate use of IVF leads to shorter time to pregnancy than strategies where IVF is used only after a trial of agents such as clomiphene or gonadotropins has led to a change in guidelines, which now suggest that the cumulative exposure to gonadotropins during the course of treatment is likely to decrease compared to earlier cohorts of women, reducing any long-term risks.
In addition, there may be cohort effects in terms of other exposures that may affect the absolute risk of some outcomes (e.g., changes in the use of postmenopausal hormone replacement therapy or ages of mammography screening affecting breast cancer risk), which in turn would impact any additional absolute risk due to exposure to infertility treatments. Because of this phenomenon, there is likely to always be some unresolvable uncertainty about long-term outcomes for both parents undergoing current infertility treatments and their children.

**Research Recommendations**

In an era of constrained resources, future clinical research, especially comparative effectiveness research—which helps resolve current uncertainties regarding clinical or policy decisions—should receive priority. For most of the KQs, there are multiple areas of remaining uncertainty based on the existing evidence. In part because of the diversity of causes and treatment options, it is difficult to make specific recommendations for specific topics.

Before setting a specific agenda for future research in infertility, we believe a more general approach to identifying priorities would be helpful. Achieving consensus on the relative priority of specific outcomes, incorporating the perspective of multiple stakeholders (similar to the approach used for developing a research agenda for comparative effectiveness research for uterine fibroids.214,215 Ideally, these outcome priorities would be used for subsequent evidence syntheses and guideline development.

As part of this consensus process, additional areas of discussion include:

- Formal consideration of the limits of acceptability for specific quantitative harms (e.g., preterm birth) and clinically meaningful differences in benefits (e.g., live birth).
- Formal discussion of the potential role of cost-effectiveness in decision making, including issues of willingness-to-pay and appropriate choice of outcome. This is particularly important because there are significant methodological challenges to the use of “standard” measures such as quality-adjusted life expectancy in the setting of infertility treatment.
- Issues related to study design, particularly from the patient stakeholder perspective. For example, in settings where patients and/or clinicians may have strong preferences for specific treatments, recruitment into RCTs may be difficult.216 In the uterine fibroid consensus process, patient stakeholders strongly preferred observational designs to randomized treatment assignment.214 Discussion of potential trade-offs between risk of bias, efficiency, ability to measure all relevant potential confounders and effect modifiers, appropriateness of alternative approaches such as Zelen randomization (where subjects are randomized prior to consent, then allowed to either receive the assigned treatment or choose the alternative217), and the likelihood that a specific study design would resolve a specific area of uncertainty should all be included.
- Issues related to data reporting. Particularly for ART and other treatments which are used for multiple indications, reporting of results separately by indication in both randomized trials and large observational studies would be extremely useful. Although these subgroup results may have insufficient power to detect clinically relevant differences within the context of individual studies (particularly RCTs), their routine publication would eventually allow synthesis of results using methods such as meta-analysis (including individual-level meta-analysis.)
The Society for Assisted Reproductive Technology Clinic Outcome Reporting System (SART CORS) and the National ART Surveillance System (NASS), which includes data submitted through SART CORS (the majority of clinics providing ART as well as a smaller number of non-SART participating clinics who report directly to the Centers for Disease Control and Prevention (CDC), are outstanding examples of what a large-scale, population-based registry can achieve in terms of providing data on treatment outcomes. However, the major limitation of the database in the past has been that data are only published on a per-cycle, rather than per-couple, basis. Recently the database methods have changed and now they are publicly reporting the cumulative success rate per patient. Results, however, are still reported at the clinic level, so patients who receive care at more than one clinic do not have the full range of outcomes captured, and there is no mechanism for prospectively collecting long-term outcomes of patients or children. Facilitating reporting of results so that outcomes are reported on a per-couple basis will substantially improve the ability to generate estimates of the likely outcome of specific ART-related decisions.

Based on input from key informants and our technical expert panel (TEP), we structured the review based on infertility diagnosis, and required studies to report outcomes specifically by diagnosis, or to adjust for diagnosis in multivariable analyses. As noted above, this led to exclusion of a number of papers, particularly those related to ART methods. There is clear evidence that the probability of some outcomes of interest, both short-term (e.g., OHSS) and long-term (certain cancers) differs based on underlying diagnosis. Although this may not be the case for all outcomes, we believe it would be helpful for future studies of interventions performed in patients with different underlying diagnoses to report results separately by diagnosis. Within an individual study powered on the basis of the total patients, estimates of diagnosis-specific outcomes may be too imprecise to confidently rule out clinically relevant differences—consistency of reporting would allow formal synthesis of estimates across studies.

We found very limited evidence on outcomes among sperm or oocyte donors. Oocyte donors, who undergo controlled ovarian hyperstimulation and oocyte retrieval in the same manner as patients undergoing IVF using their own eggs, have, in theory, at least the same risk of short-term adverse events as patients. The frequency with which oocyte donors are used is increasing, and evidence from the SART CORS database suggests that the risk of certain pregnancy complications is lower when donor oocytes are used. If demand for donor oocytes continues to increase, much more evidence on the specific short- and long-term outcomes of donation (especially if a donor undergoes multiple cycles) is needed.

Conclusions

There is evidence supporting some strategies for treatment of infertility, both for specific diagnoses and for couples with any diagnosis, in part because of recent adaptation of more rigorous methods for evaluating treatments for infertility, particularly regarding treatments for PCOS and approaches to timing of interventions in patients undergoing ART. In addition, ongoing refinements to the SART CORS database continue to make it a valuable resource, particularly for data on short-term outcomes. However, given the diversity of infertility causes and treatments, there is considerable residual uncertainty about the optimal treatment options for specific patients. Consensus on which outcomes to report (such as encouraging reporting of live birth rates on a per couple basis as well as per cycle, and, for studies of treatment such as ART, reporting of both overall and diagnosis-specific outcomes) and which areas of uncertainty are
most important to resolve (in order to prioritize research) is needed to improve the ability of patients and clinicians to make decisions about the most appropriate treatment.
References


Introduction

Background

Condition

“Infertility” has traditionally been defined as failure to achieve pregnancy after 12 months of regular unprotected intercourse with the same partner (or after 6 months for women greater than 35 years of age). However, as many as half of such couples will conceive without intervention over the next 12–24 months. Because of this, the term “subfertility” is preferred by many.1 From a population perspective, couples who meet the dichotomous criteria for “infertility” include couples who are “normal” but who are in the upper end of the population distribution for “time to pregnancy”, and couples who have a physiological or anatomical cause for a prolonged time to pregnancy. However, to be concise, we use the term “infertility” throughout this report.

Self-reported infertility in the United States, using the 12-month definition, affected approximately 6 percent of married women aged 15–44 in the 2006-2010 National Survey of Family Growth (the most recent available data).2 In one population-based study, approximately 10 percent of pregnant women reported receiving infertility treatment, with 29 percent of these women using fertility-enhancing medications; 21 percent using assisted reproductive technology (ART), including in vitro fertilization (IVF); 15 percent using artificial insemination with fertility-enhancing drugs; and 23 percent using other treatments, including surgery.3 Other estimates of the prevalence of infertility treatment are similar.4-8 Particularly in the US, where availability of infertility services is variable depending on a number of factors, particularly insurance coverage, utilization of infertility treatments may underestimate the overall burden of infertility.

The most common demographic factor associated with female infertility is “advanced reproductive age”; although the probability of pregnancy begins to decline by the mid-20’s, the slope of decline sharply increases by age 35.9 For example, the prevalence of “unexplained infertility” (infertility with no other documented female or male diagnosis) is substantially higher in older women,10 and “diminished ovarian reserve,” which is most commonly associated with increased age, is the single most common diagnosis among women undergoing ART, accounting for 27.5 percent of cycles.11 Other common causes of female infertility include polycystic ovary syndrome (PCOS), endometriosis, and occlusion of the fallopian tubes from prior infectious disease.6 A growing number of women also experience infertility secondary to cancer treatment.12-14. Although there are other potentially treatable causes of infertility (conditions other than PCOS which affect ovulatory function, congenital uterine anomalies, uterine fibroids), this review is based on the most common conditions, based on input from our Technical Expert Panel.

Based on estimates of patients attending ART clinics, isolated male factor infertility affects approximately 17 percent of couples seeking treatment, with 34.6 percent of couples having both male and female diagnoses.15

Treatment Strategies

Treatment options are usually dependent on the underlying etiology of infertility. For female causes, options include surgical management of tubal occlusion, surgical treatment of endometriosis, ovarian “drilling” for treatment of PCOS, use of ovulation-induction agents
including oral (clomiphene citrate or letrozole) and injected drugs (gonadotropins), artificial insemination with either partner or donor sperm (depending on partner fertility status), and ART, which includes both traditional IVF (fertilization of the egg by the sperm occurs without direct manipulation) and IVF with intra-cytoplasmic sperm injection (ICSI) (fertilization occurs via direct injection of sperm into the egg). Treatment options for male factor infertility include medical treatment of a diagnosed endocrinopathy or other conditions affecting sperm production, empiric treatments with hormonal or other agents, surgical management of varicocele, intrauterine insemination, IVF, ICSI, or use of donor sperm. Options appropriate for some diagnoses (e.g., ovulation induction in PCOS or unexplained infertility) may not be appropriate for others (e.g., women with documented tubal occlusion). In other cases, the appropriate comparisons may involve sequencing or combinations of treatment options—for example, one strategy might consist of several cycles of ovulation induction, followed by ART only if pregnancy does not occur, compared to proceeding directly to ART. Note that throughout this report, we use the term “adjunct treatments” to refer to interventions performed within a major treatment category (for example, comparison of metformin to placebo as pretreatment in women with PCOS undergoing IVF).

Benefits

There has been ongoing debate about the most appropriate outcome for evaluation of infertility treatments—ovulation (in anovulatory women such as PCOS patients), pregnancy, live birth, or term live birth. However, there is a growing consensus that live birth is the most important patient-centered outcome. Trade-offs in outcomes (particularly multiple gestations), time to pregnancy, and out-of-pocket costs might be different among the various treatment strategies even if cumulative live birth rates are identical.

Harms

Different treatments also carry different safety risks. There are known short-term risks such as ovarian hyperstimulation syndrome (OHSS) or acute risks associated with any surgery. Surgery may have additional longer-term risks which may affect subsequent fertility (such as scarring or decreased ovarian reserve with procedures such as laparoscopic ovarian drilling (LOD). The literature suggests that observed associations between infertility treatment and female reproductive cancers, particularly ovarian cancer, are likely the result of the underlying infertility rather than treatment itself. There is, however, some uncertainty surrounding some cancer outcomes in subgroups of patients.

Some adverse pregnancy outcomes, such as preterm birth, are associated with infertility treatment; however, many of the conditions associated with infertility are also associated with these adverse outcomes, complicating assessment of comparative effectiveness. There may also be treatments that have unclear effects on the long-term health in children born following these treatments For example, there is the possibility that epigenetic changes from treatments such as IVF/ICSI may lead to increased risk of some disorders (e.g., Beckwith-Wiedemann syndrome) later in life—or that an increase in multiple births or other causes of prematurity or fetal growth restriction/low birthweight from treatments may result in poor neurodevelopmental outcomes. Finally, infertility clearly has an emotional impact and the comparative effects of infertility treatments on quality of life are an important consideration for both women and men.
There may be significant variation in outcomes of different treatments in specific subpopulations. For example, age affects the likelihood of conception, and the risk of many pregnancy complications associated with infertility treatments, such as preterm birth or low birthweight, are also increased with higher maternal age. Obesity is common in women with PCOS, and, like older maternal age, is also associated with adverse pregnancy outcomes independent of its association with infertility. The utilization and outcomes of infertility treatment differ among different racial and ethnic groups, even after adjusting for insurance coverage.33-36

Finally, a unique subpopulation is women who donate oocytes for use by other couples in ART. An increasing number of women undergoing ART are receiving donor oocytes,37 and there are almost no data on the long-term safety of multiple courses of ovulation induction for the purposes of oocyte donation.38 In addition, there are complex ethical and legal considerations, including the balance between fair compensation and inducement,39 and sharing information about donors with recipients.40

Complexity of Decision Making for Treatments of Infertility

Infertility treatment is a topic where decision making is particularly complex for patients, clinicians, and policymakers. Decision making involves both partners (although the intensity and risks of treatment are quite different), consideration of outcomes for both parents and infants over short- and long-term time frames, trade-offs between short-term success and long-term adverse outcomes, and in some cases preferences for process as well as outcome. In addition, time is an important consideration, particularly for women aged 35 and older. There is clear variation in patient preferences for different treatments and outcomes, and there has been relatively little empirical work focused on the decision-making aspects of infertility treatment. There are large differences in the costs of different infertility treatments and variation in the degree of coverage for infertility diagnosis and treatment, and many patients face significant out-of-pocket costs.41 There is substantial evidence that the availability of coverage affects access to treatment and treatment choices.42-45 Time lost from work may also be a consideration (particularly in the context of the need to make out-of-pocket payments).

There are a number of areas where controversy or uncertainty about the evidence adds to the difficulty of decision making. For example, the optimal trade-off between ART success and the risk of preterm birth and long-term health outcomes (such as neurodevelopmental problems) in infants associated with the number of embryos transferred is unclear. All things being equal, transfer of more embryos results in both a greater chance of success in a given ART cycle and a greater chance of multiple pregnancies—single-embryo transfer greatly reduces the chance of multiple gestation, but may require more cycles to achieve a pregnancy.46,47 Other areas of uncertainty include optimal timing of embryo transfer48 and use of fresh versus frozen embryos,49,50 in terms of both achieving pregnancy and outcomes of those pregnancies, as well as timing of ART relative to other options, especially since the risk of higher order multiples (triplets or higher) is greater with ovulation induction, although ART is more invasive and expensive on a per-cycle basis.51-53

Limitations of the Evidence Base

Methodological limitations of the literature contribute to the uncertainty. For example, the National ART Surveillance System (NASS) is an excellent resource for observational data on U.S. population-based outcomes for ART. However, it has been limited by (a) use of the ART
cycle (rather than the individual patient) as the unit of analysis; (b) lack of long-term follow-up data for individual patients; and (c) some concern about underreporting of some adverse outcomes. NASS 2.0, introduced at the end of 2016, now includes both unique patient and cycle identifiers, meaning that cumulative success rates per patient should be available in future years. In addition, mechanisms for capturing outcomes from patients who receive care at multiple clinics have been put into place. On the other hand, randomized controlled trials (RCTs) may not provide data on important long-term outcomes, or may be underpowered to detect clinically relevant differences in complications of treatment.

Scope and Key Questions

Scope of the Review

The present review evaluates the comparative safety and effectiveness of available treatment strategies for women of reproductive age (18–44) who are infertile due to PCOS, endometriosis, unknown reasons, or tubal or peritoneal factors; the comparative safety and effectiveness of available treatment strategies for couples with male factor infertility; and the short- and long-term health outcomes of both oocyte and sperm donors. For all questions, we consider only treatment options begun after completion of a diagnostic evaluation.

Key Questions

The specific Key Questions (KQs) addressed in this review are listed below, and Figure 1 displays the analytic framework that guided our work.

- **KQ 1.** What are the comparative safety and effectiveness of available treatment strategies for women with **polycystic ovary syndrome (PCOS)** who are infertile and who wish to become pregnant?
  - KQ 1a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, body mass index (BMI), presence of other potential causes of female infertility, or presence of male factor infertility?

- **KQ 2.** What are the comparative safety and effectiveness of available treatment strategies for women with **endometriosis** who are infertile and who wish to become pregnant?
  - KQ 2a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, stage of endometriosis, presence of other potential causes of female infertility, or presence of male factor infertility?

- **KQ 3.** What are the comparative safety and effectiveness of available treatment strategies for women who are infertile for **unknown reasons** and who wish to become pregnant?
  - KQ 3a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI,
presence of other potential causes of female infertility, or presence of male factor infertility?

- **KQ 4.** What are the comparative safety and effectiveness of available treatment strategies for women with tubal or peritoneal factors (e.g., pelvic adhesions) who are infertile and who wish to become pregnant?
  - **KQ 4a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, presence of other potential causes of female infertility, or presence of male factor infertility?

- **KQ 5.** What are the comparative safety and effectiveness of available treatment strategies for couples with male factor infertility and no evidence of an underlying diagnosis associated with infertility in the female partner?
  - **KQ 5a.** Does the optimal treatment strategy vary by characteristics in either partner such as age, ovarian reserve, race, or BMI?

- **KQ 6.** What are the short- and long-term health outcomes of donors in infertility?
  - **KQ 6a.** For female oocyte donors:
    3. Do specific aspects of the pre-donation evaluation identify potential donors at greater risk for short- or long-term adverse outcomes (e.g., OHSS, quality-of-life issues)?
    4. Do short- and long-term outcomes differ among different stimulation/retrieval protocols?
  - **KQ 6b.** For male semen donors:
    1. Are there long-term health, quality-of-life, or other adverse outcomes associated with donation?
Abbreviations: ART=assisted reproductive technology; BMI=body mass index; GnRH=gonadotropin-releasing hormone; KQ=Key Question; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome
Figure 1 depicts the KQs within the context of the populations, interventions, comparators, outcomes, timings, and settings (PICOTS) considered in this review. The figure illustrates how a wide range of treatments for infertility may result in intermediate outcomes such as time to pregnancy and/or final outcomes such as live birth (single or multiple) or costs in couples with different underlying causes of infertility. A separate KQ focuses on outcomes in female and male donors in infertility. Short- and long-term adverse effects may occur at any point during treatment and may affect donors, patients, and/or children. Optimal treatment strategies may vary by important patient characteristics and/or by setting/provider.

Organization of This Report

The remainder of the report details our methodology and presents the results of our literature synthesis, with summary tables and strength of evidence grading for major comparisons and outcomes. In the discussion section, we offer our conclusions, summarized findings, and other information that may be relevant to translating this work for clinical practice and future research.

Appendixes provide further details on our methods and the studies we assessed, as follows:

- Appendix A. Exact Search Strings
- Appendix B. Data Abstraction Elements
- Appendix C. List of Included Studies
- Appendix D. List of Excluded Studies
- Appendix E. Characteristics of Included Studies
- Appendix F. AMSTAR Quality Assessment for Systematic Reviews
- Appendix G. Risk of Bias Assessment for Included Studies
- Appendix H. Supplemental Project To Assess the Transparency of Reporting for Trials Evaluating Treatment for Infertility

A list of abbreviations and acronyms is provided at the end of the report.
Methods

Methods for this systematic review follow the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist. See the review protocol (http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=2131&pageaction=displayproduct) for full details.

Topic Refinement and Review Protocol

The topic of this report and preliminary Key Questions (KQs) arose through a public nomination and initial development by the Scientific Resource Center (SRC) for AHRQ’s Effective Health Care (EHC) program. During the subsequent topic refinement phase, a panel of key informants gave input to the Evidence-based Practice Center (EPC) on the KQs to be examined; these KQs were posted on AHRQ’s EHC website for public comment in June 2015 for 3 weeks and revised in response to comments. We then drafted a protocol for the systematic review and recruited a Technical Expert Panel (TEP) to provide high-level content and methodological expertise throughout the development of the review. The Key Informants and TEP represented members of medical professional societies and clinician/researchers in the areas of obstetrics and gynecology, assisted reproductive technology, and reproductive medicine; scientific experts; payers; Federal agencies; and patients/consumers. The finalized protocol is posted on the EHC website. The PROSPERO registration is CRD42016025750.

Literature Search Strategy

Search Strategy

To identify relevant published literature, we searched PubMed®, Embase®, and the Cochrane Database of Systematic Reviews, limiting the searches to studies conducted in adults and published from January 1, 2007, to October 3, 2018. We selected the 2007 start date to establish a one-year overlap with the search dates from a 2008 AHRQ evidence report that assessed assisted reproductive technology (ART). Also, the key informants felt that existing Cochrane reviews would identify older relevant high-quality studies, particularly evidence from randomized controlled trials (RCTs), while primary studies and other systematic reviews published after 2008 would identify studies most relevant to current practice in infertility. An experienced search librarian guided all searches. The exact search strings used are given in Appendix A.

We supplemented the electronic searches with a manual search. The reference lists for identified pivotal articles were manually searched and cross-referenced against our database, and additional relevant articles not already under consideration were retrieved for screening. All citations were imported into an electronic bibliographical database (EndNote® Version X8; Thomson Reuters, Philadelphia, PA).

To ascertain publication bias in recent studies, we searched ClinicalTrials.gov to identify completed but unpublished studies. We also explored publication bias in any quantitative synthesis of the included literature through meta-analysis techniques. In Appendix H, we use this report to explore in more detail the utility of ClinicalTrials.gov for detecting selective reporting, and the impact of selective reporting on the estimates of treatment effect. Note that these evaluations of publication bias require studies to have been registered on ClinicalTrials.gov or
published in the peer-reviewed literature and therefore are not able to reflect studies which found no difference or negative results but did not reach one of these outlets.

Approaches to identifying relevant gray literature included notification to stakeholders of the opportunity to submit scientific information packets of material relevant to the KQs. This notification was coordinated by the SRC. We also searched the ClinicalTrials.gov study registry and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) search portal to identify potentially relevant study records, and subsequently searched for relevant articles from completed studies.

Inclusion and Exclusion Criteria

We specified our inclusion and exclusion criteria based on the PICOTS (populations, interventions, comparators, outcomes, timing, and settings) identified for each question. Note that the outcomes of interest are ordered in approximate relative importance to patients, based on input from topical experts and Key Informants, rather than temporal occurrence in the clinical pathway. Table 1 lists inclusion and exclusion criteria.

Table 1. Inclusion and exclusion criteria

<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
</table>
| Populations    |                    | Individuals younger than 18 or 45 and older
<p>| KQs 1-4:       | Women of reproductive age (18-44) with no pregnancy after 12 months of regular intercourse for women under 35, or 6 months for women 35 and older (studies which use alternate definitions of infertility based on different age or duration criteria may be included if justified), and | Study does not report outcomes of interest by underlying diagnosis or by using a multivariate model that includes diagnosis as one of the covariates |
|                | - Diagnosed PCOS (KQ 1) | |
|                | - Diagnosed endometriosis (KQ 2) | |
|                | - No other diagnosed cause of infertility (KQ 3) | |
|                | - Identified tubal or peritoneal disease potentially amenable to surgical interventions (hydrosalpinx, unilateral occlusion, prior tubal sterilization) (KQ 4) | |
|                | Subpopulations of interest include groups differing in: | |
|                | - Age; race/ethnicity; obesity/BMI; ovarian reserve; history of prior treatments; primary vs. secondary infertility; maternal parity; insurance status (KQs 1-4) | |
|                | - Diagnostic criteria/evaluation; presence or absence of male factor infertility, other female causes of infertility, or common comorbidities such as hypertension and diabetes (KQ 1) | |
|                | - Diagnostic criteria/evaluation; stage of endometriosis; presence or absence of male factor infertility, presence of endometrioma, other female causes of infertility, or common comorbidities such as hypertension and diabetes (KQ 2) | |
|                | - Diagnostic criteria/evaluation; presence or absence of common comorbidities such as hypertension and diabetes; women without male partners (single women or lesbian couples) (KQ 3) | |
|                | - Anatomic cause of tubal occlusion (e.g., prior sterilization vs. adhesions) (KQ 4) | |
| KQ 5:          | Couples which include men partnered with women of reproductive age (as defined in other KQs), with no documented female cause of infertility and documented male infertility. | |
|                | Subpopulations of interest include groups differing by cause of male infertility (identified hormonal cause, varicocele, idiopathic), age | |</p>
<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>(male and female), race/ethnicity, obesity/BMI, history of prior treatments, primary vs. secondary infertility, diagnostic criteria used for male infertility, insurance status, and presence or absence of common comorbidities such as hypertension and diabetes.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ 6:</td>
<td>Women of reproductive age (18-44) who are potential donors of oocytes for ART, and males donating semen for intrauterine insemination or ART</td>
<td></td>
</tr>
</tbody>
</table>

### Interventions

| KQ 1: Clomiphene citrate, letrozole, diet/exercise/other weight loss strategies, timed intercourse using various technologies in conjunction with oral ovulation induction, metformin, combination oral medications, ovulation induction with gonadotropins with or without intrauterine insemination (IUI), surgery (ovarian drilling), ART (IVF and ICSI) with patient and donor oocytes | | |
| Surgical excision of endometriotic implants, alternative surgical approaches to destruction of lesions (e.g., laser vaporization), gonadotropin-releasing hormone agonists or antagonists, timed intercourse with various technologies, ovulation induction with gonadotropins with or without IUI, ART (IVF and ICSI) with patient and donor oocytes | |
| Timed intercourse with various technologies, oral ovulation induction agents (e.g., clomiphene citrate), superovulation with gonadotropins with and without IUI, ART (IVF and ICSI) with patient and donor oocytes, watchful waiting | |
| Surgical repair, ART (IVF and ICSI) with patient and donor oocytes | |
| ICSI (note that interventions and comparators may vary depending on underlying cause of male factor infertility), testicular sperm extraction, vasectomy reversal, surgical repair of varicocele, IUI, donor insemination, ART, treatment of underlying endocrinopathy | |
| Pre-donation testing strategies; ovulation induction with gonadotropins using different induction/retrieval protocols; semen donation (men) | |

### Comparators

| KQ 1: Any other active intervention (e.g., clomiphene vs. metformin), or timing/sequence of interventions (e.g., ovulation induction/IUI followed by ART if unsuccessful vs. proceeding directly to ART, or timed intercourse with oral medications or injectable gonadotropins) | | |
| 2: Either direct between two alternatives (e.g., surgery vs. GnRH agonists/antagonists), or timing/sequence of interventions (e.g., ovulation induction/IUI followed by ART if unsuccessful vs. proceeding directly to ART) | | |
| KQ 3: Any other active intervention, or timing/sequencing of timing/sequence of interventions (e.g., ovulation induction/IUI followed by ART if unsuccessful vs. proceeding directly to ART) | | |
| KQ 4: Other active interventions (including combinations of therapy such as surgical removal of hydrosalpinx followed by ART) | | |
| KQ 5: Other active interventions | | |
| KQ 6 (women): Pre-donation testing strategies; controlled ovarian hyperstimulation with gonadotropins using different induction/retrieval protocols; non-donors (women and men) | | |

### Outcomes

<p>| KQ 1-5: | | |
| Live birth (both cumulative and per cycle) | | |
| o Live singleton birth | | |
| o Live multiple birth | | |
| Pregnancy complications | | |
| o Multiple births (and associated complications) | | |</p>
<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ectopic pregnancies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miscarriage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Death</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birthweight (categorized as low birthweight/normal birthweight)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Congenital anomalies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time to pregnancy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Calendar time (months)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health system</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Societal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term adverse effects of treatments</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OHSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Surgical complications</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term outcomes (child)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurodevelopmental/other issues related to prematurity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific issues related to infertility treatment (epigenetic changes, sex chromosomal abnormalities, etc.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer (all types)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-term outcomes (maternal)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subsequent fertility</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

KQ 6:

- **Women:**
  - Short-term adverse effects of treatments
    - OHSS
    - Surgical complications
    - Adverse effects of treatments
  - Long-term outcomes (donor)
    - Downstream fertility
    - Cancer
    - Age at menopause
  - Quality-of-life outcomes

- **Men:**
  - Quality-of-life outcomes
  - Short- and long-term health outcomes

### Timing

KQs 1-5:

- **Short-term**
  - From beginning of treatment through first 12 months of life if live birth occurs
- **Long-term**
  - 12 months or more from completion of treatment (no live birth) or from date of live birth

KQ 6:

- **Short-term:**
  - From time of beginning donation process to 12 months after donation
- **Long-term:**
  - 12 months or more from time of first donation

### Settings

- Subspecialty practice (infertility specialist) (KQs 1-6)
- General gynecology practice (KQs 1-5)
- Family practice/general internist/nurse practitioner/other non-gynecologist primary care provider (KQs 1-6)
- Male reproductive medicine specialist/urologist (KQ 5)
<table>
<thead>
<tr>
<th>PICOTS Element</th>
<th>Inclusion Criteria</th>
<th>Exclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td>• Original data&lt;br&gt;• RCTs, prospective and retrospective observational studies with comparator; for test characteristics, cross-sectional studies were acceptable if they included patients with diagnostic uncertainty and direct comparison of test results with an appropriate reference standard&lt;br&gt;• Study design limitations by outcome type:&lt;br&gt;  o KQs 1-5:&lt;br&gt;     ▪ Short-term outcomes: RCTs of any sample size; observational studies of ≥100 subjects presenting data from the National ART Surveillance System&lt;br&gt;     ▪ Long-term outcomes: RCTs of any sample size; observational studies of ≥100 subjects&lt;br&gt;  o KQ 6 (all outcomes):&lt;br&gt;     ▪ RCTs of any sample size; observational studies of ≥100 subjects</td>
<td>Editorials, nonsystematic reviews, abstracts only, letters, case series, case reports, articles that have been retracted or withdrawn</td>
</tr>
<tr>
<td>Publications</td>
<td>• English-language only&lt;br&gt;• Published January 1, 2007, to present&lt;br&gt;• Relevant systematic reviews, meta-analyses, or methods articles</td>
<td>Non-English language articles(^5)</td>
</tr>
</tbody>
</table>

\(^{a}\)Studies which included patients outside this age range were included if findings were reported separately for this age group of interest or if at least 80% of women were within this age range. \(^{b}\)Non-English articles were excluded due to: (1) the high volume of literature available in English-language publications (including the majority of known important studies); and (2) concerns about the applicability of non-English publication studies to populations in the United States.

Abbreviations: ART=assisted reproductive technology; BMI=body mass index; GnRH=gonadotropin-releasing hormone; hCG=human chorionic gonadotropin; ICSI=intracytoplasmic sperm injection; IUI=intruterine insemination; IVF=in vitro fertilization; KQ=Key Question; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome; PICOTS=Populations, Interventions, Comparators, Outcomes, Timing, Settings; RCTs=randomized controlled trials

**Study Selection**

For citations retrieved from PubMed, Embase, and the Cochrane Database of Systematic Reviews, two reviewers independently screened each title and abstract for potential relevance to the research questions using prespecified inclusion/exclusion criteria described in Table 1. Articles included by either reviewer underwent full-text screening.

At the full-text screening stage, two reviewers independently reviewed the full text of each article and indicated a decision to include or exclude the article for data abstraction. When paired reviewers arrived at different decisions about whether to include or exclude an article, or about the reason for exclusion, we reconciled the difference through review and discussion among investigators. Articles meeting eligibility criteria were included for data abstraction. All screening results were tracked using the DistillerSR data synthesis software program (Evidence Partners Inc., Manotick, ON, Canada).

Appendix C provides a list of all articles included for data abstraction. Appendix D provides a list of articles excluded at the full-text screening stage, with reasons for exclusion.

While systematic reviews and meta-analyses were not study designs qualifying for inclusion and abstraction under our screening criteria, we did flag relevant articles of these types as part of the screening process. Component references from these systematic reviews were reviewed and when studies met our inclusion criteria, they were included in our report. For systematic reviews which were identified as relevant to the individual KQs but included mostly studies prior to 2007, we summarize the findings from these existing reviews and the consistency of their findings with those from our included studies in the appropriate results sections.
Data Extraction

The investigative team created data abstraction forms that were programmed in the DistillerSR software for collection of data from included studies. The abstraction forms were pilot-tested with a sample of included articles to ensure that all relevant data elements were captured and that there was consistency and reproducibility between abstractors. Based on their clinical and methodological expertise, a pair of researchers were assigned to abstract data from each of the eligible articles. One researcher abstracted the data, and the second over-read the article and the accompanying abstraction to check for accuracy and completeness. Disagreements were resolved by consensus or by obtaining a third reviewer’s opinion if consensus could not be reached.

We designed the data abstraction forms to collect the data required to evaluate the specified eligibility criteria for inclusion in this review, as well as demographic and other data needed for determining outcomes (intermediate, final, and adverse events outcomes). We paid particular attention to describing the details of the treatment (e.g., for comparisons of in vitro fertilization [IVF] to other therapies, the specific IVF protocol used), patient characteristics (e.g., age of female partners, presence or absence of male factor infertility), setting (e.g., U.S.- vs. non-U.S.-based studies), and study design (e.g., RCT vs. observational) that may be related to outcomes. In addition, we described comparators carefully, as treatment standards may have changed during the period covered by the review. The safety outcomes were framed to help identify adverse events, including those from medical therapies (e.g., ovarian hyperstimulation syndrome) and those resulting from procedural complications. Data necessary for assessing quality and applicability, as described in AHRQ’s *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*, were also abstracted. A complete list of data abstraction elements is provided in Appendix B.

Quality (Risk of Bias) Assessment of Individual Studies

We assessed methodological quality, or risk of bias, for each individual study using a components approach, assessing each study for specific aspects of design or conduct (such as allocation concealment for RCTs, or use of methods to address potential confounding), as detailed in AHRQ’s *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Briefly, we rated each study as being of good, fair, or poor quality based on its adherence to well-accepted standard methodologies. For each study, one investigator assigned a summary quality rating, which was then reviewed by a second investigator; disagreements were resolved by consensus or by a third investigator if agreement could not be reached. Table 2 describes the overall study quality assessment ratings. Appendix G presents the risk of bias assessment components for the individual included studies.
Table 2. Definitions of overall quality ratings

<table>
<thead>
<tr>
<th>Quality Rating</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good (low risk of bias)</td>
<td>These studies had the least bias, and the results were considered valid. These studies adhered to the commonly held concepts of high quality, including the following: a clear description of the population, setting, approaches, and comparison groups; appropriate measurement of outcomes; appropriate statistical and analytical methods and reporting; no reporting errors; a low dropout rate; and clear reporting of dropouts.</td>
</tr>
<tr>
<td>Fair (moderate risk of bias)</td>
<td>These studies were susceptible to some bias, but not enough to invalidate the results. They did not meet all the criteria required for a rating of good quality because they had some deficiencies, but no flaw was likely to cause major bias. The study may have been missing information, making it difficult to assess limitations and potential problems.</td>
</tr>
<tr>
<td>Poor (high risk of bias)</td>
<td>These studies had significant flaws that might have invalidated the results. They had serious errors in design, analysis, or reporting; large amounts of missing information; or discrepancies in reporting.</td>
</tr>
</tbody>
</table>

The grading was outcome-specific such that a given study that analyzed its primary outcome well but did an incomplete analysis of a secondary outcome could be assigned a different quality grade for each of the two outcomes. Studies of different designs were graded within the context of their respective designs as good, fair, or poor (Appendix G).

We also rated the quality of systematic reviews that were identified and discussed in the report using AMSTAR (A Measurement Tool to Assess the Methodological Quality of Systematic Reviews). For each study, one investigator assigned a summary quality rating, a second investigator reviewed the rating; disagreements were resolved by consensus or by a third investigator. Reviews were then assigned overall quality scores according to the following categories:

- Good (low risk of bias)—Systematic reviews that have few or no methodological shortcomings and a low risk of bias.
- Fair (moderate risk of bias)—Systematic reviews that have some methodological flaws but the investigators conclude that the flaws will not seriously bias or invalidate the results.
- Poor (high risk of bias)—Systematic reviews that contain a serious flaw or flaws that, in the judgment of the investigators, are highly likely to bias or invalidate the results.

The AMSTAR quality assessment components for the individual systematic reviews are detailed in Appendix F.

Data Synthesis

We began by summarizing key features of the included studies for each KQ. To the degree that data are available, we abstracted information on study design; patient characteristics; clinical settings; interventions; and intermediate, final, and adverse event outcomes. If not reported, 95-percent confidence intervals for dichotomous outcomes (e.g., live birth rates) were calculated from the numbers provided in the study, in order to characterize the degree of precision of a particular estimate. This helped inform grading of the strength of evidence, as well as provided insight about the degree to which lack of statistical power may have affected study conclusions about lack of a treatment effect.
We then determined the feasibility of completing a quantitative synthesis (i.e., meta-analysis, decision analysis, or simulation model). For a meta-analysis, feasibility depends on the volume of relevant literature (requiring at least three relevant studies), conceptual homogeneity of the studies (similar intervention comparisons and outcome definitions), completeness of the reporting of results, and the adequacy and completeness of any existing meta-analyses.

When the above criteria were met and a meta-analysis was considered appropriate, we used random-effects models within the Comprehensive Meta-Analysis software to synthesize the available evidence quantitatively. We tested for heterogeneity using graphical displays and test statistics (Q and I² statistics), while recognizing that the ability of statistical methods to detect heterogeneity may be limited. For comparison, we also performed fixed-effect meta-analyses. We present summary estimates, standard errors, and confidence intervals. We anticipated that intervention effects may be heterogeneous. We hypothesized that the methodological quality of individual studies, study type, the characteristics of the comparator, and patients’ underlying clinical presentation would be associated with the intervention effects. If there were sufficient studies, we performed subgroup analyses and/or meta-regression analyses to examine these hypotheses. We performed quantitative and qualitative syntheses separately by study type and discuss their consistency qualitatively.

**Strength of the Body of Evidence**

We rated strength of evidence using the approach described in the AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews.56,115,116 We graded the strength of evidence separately for each outcome; thus, the strength of evidence for two separate outcomes in a given study may be graded differently. The grades are presented in the strength of evidence tables throughout the report.

Briefly, the approach requires assessment of five domains: study quality (previously named risk of bias, and described above), consistency, directness, precision, and reporting bias (which includes publication bias, outcome reporting, and analysis reporting biases). When the body of evidence for a particular outcome included both RCTs and observational studies, we graded each study type separately using design-specific criteria. In considering the overall strength of the entire body of evidence, we considered the extent to which the observational evidence was consistent with RCT data, particularly with regard to direction and magnitude of effect. We also explored the consistency of our findings with recent systematic reviews and discussed agreement or disagreement, along with possible causes for disagreement and impact on strength of evidence ratings, in the text. Because of the risk of unmeasured confounding, observational studies generally would not contribute to estimates of the magnitude or precision of effect, when RCT data were available. If there were other issues (such as differences in when and where RCTs were performed compared to observational studies, and how these differences might affect applicability), this would generally increase uncertainty about the magnitude and precision of any treatment effect.117 The five domains were considered qualitatively, and a summary rating of high, moderate, or low strength of evidence was assigned after discussion by two reviewers. When no evidence was available or when evidence on the outcome was too weak, sparse, or inconsistent to permit any conclusion to be drawn, a grade of “insufficient” was assigned. This four-level rating scale consisted of the following definitions:

- **High Strength of Evidence**—We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable; i.e., another study would not change the conclusions.
- Moderate Strength of Evidence—We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low Strength of Evidence—We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.
- Insufficient Strength of Evidence—We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome.

**Applicability**

We assessed applicability across our KQs using the method described in AHRQ’s *EPC Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. In brief, this method uses the PICOTS format as a way to organize information relevant to applicability. The most important issue with respect to applicability is whether the outcomes were different across studies that recruit different populations (e.g., age groups, U.S. vs. non-U.S. settings) or used different methods to implement the interventions of interest; that is, important characteristics are those that affect baseline (control group) rates of events, intervention group rates of events, or both. We used a checklist applied to each abstracted study to guide the assessment of applicability (Appendix B). For each study, one investigator assigned a summary quality rating, which was then reviewed by a second investigator; disagreements were resolved by consensus or by a third investigator if agreement could not be reached. We summarize issues of applicability qualitatively.

**Peer Review and Public Commentary**

Experts in the fields of reproductive endocrinology, reproductive epidemiology, urology, and women’s reproductive health, and individuals representing stakeholder and user communities were invited to provide external peer review of the draft report. AHRQ, an associate editor, and members of the TEP were also invited to provide comments. In addition, the draft report was posted on the AHRQ EHC website for public comment from April 3, 2018, through May 1, 2018. We have addressed all reviewer comments, revising the text as appropriate, and documented everything in a disposition of comments report that will be made available 3 months after the Agency posts the final report on the EHC website. A list of peer reviewers submitting comments on the draft report is provided in the front matter of this report.
Results

We begin by describing the results of our literature searches. We then provide an overall description of the included studies. The remainder of the chapter is organized by Key Question (KQ). Under each of the six KQs, we begin with a brief description of the included studies, followed by a bulleted list of the key points of the findings and a detailed synthesis of the evidence. Within each KQ the detailed syntheses are organized first by treatment comparison and then by outcome. The outcomes of interest are ordered in approximate relative importance to patients, based on input from topical experts and Key Informants, rather than temporal occurrence in the clinical pathway: live birth, pregnancy complications, neonatal outcomes, time to pregnancy, costs, short term adverse effects of treatment, and long term outcomes. We conducted quantitative syntheses where possible, as described in the Methods chapter. Although not considered as formal included articles, we discuss findings from relevant systematic reviews – and whether these findings are consistent or not with the evidence from our included articles. We end each treatment section by highlighting any evidence for specific subgroups of interest. Each KQ results section concludes with a summary of the strength of evidence for the main findings. For findings applicable across all KQs, please refer to “Key Findings Across All Infertility Diagnoses,” which is presented at the end of the results section. For a list of abbreviations, please refer to the end of the report.

Results of Literature Searches

Searches of PubMed, Embase, and the Cochrane Database of Systematic Reviews yielded 21,467 citations, 17,263 of which were unique. Manual searching of gray literature databases and bibliographies of key articles or referral by investigators identified 128 additional citations, for a total of 17,391 citations. We received no responses from manufacturers to our requests for scientific information packets. After applying inclusion/exclusion criteria at the title-and-abstract level, 1,909 full-text articles were retrieved and screened. Of these, 1,748 were excluded at the full-text screening stage, leaving 161 articles for data abstraction. These 161 articles described 151 unique studies. Studies with more than one article are listed in Table 3. Note that although four studies used the National ART Surveillance System (NASS), they were not considered overlapping studies in terms of the underlying patient population and so are reported separately.

### Table 3. Key to primary and companion articles

<table>
<thead>
<tr>
<th>Study Designation</th>
<th>Primary Abstracted Article</th>
<th>Companion Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIFEStyle</td>
<td>Mutsaerts, 2016</td>
<td>van Oers, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>van Oers, 2018</td>
</tr>
<tr>
<td>NASS (National Assisted Reproductive Technology Surveillance System)</td>
<td>Butts, 2014122</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Kawwass, 2013</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Luke, 2010</td>
<td>None</td>
</tr>
<tr>
<td></td>
<td>Nangia, 2011</td>
<td>None</td>
</tr>
<tr>
<td>OMEGA Project</td>
<td>Spaan, 2015</td>
<td>van Leeuwen, 2011</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Spaan, 2016</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legro, 2006</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legro, 2014</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Legro, 2012</td>
</tr>
<tr>
<td>SUIT (Scottish Unexplained Infertility Trial)</td>
<td>Bhattacharya, 2008</td>
<td>None</td>
</tr>
</tbody>
</table>
Figure 2 depicts the flow of articles through the literature search and screening process.

**Figure 2. Literature flow diagram**

<table>
<thead>
<tr>
<th>Study Designation</th>
<th>Primary Abstracted Article</th>
<th>Companion Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>Brinton, 2015(^{136})</td>
<td>Brinton, 2014(^{137})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brinton, 2013(^{138})</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Trabert, 2013(^{139})</td>
</tr>
<tr>
<td>None</td>
<td>Custers, 2012(^{139})</td>
<td>Steures, 2006(^{140a})</td>
</tr>
<tr>
<td>None</td>
<td>Nahuis, 2011(^{141})</td>
<td>Nahuis, 2012(^{142})</td>
</tr>
</tbody>
</table>

\(^{a}\)Used for background information only.


Duplicates removed: 4,204

Citations identified through gray lit/manual searching or referral by investigators: 128

15,482 abstracts excluded

1,748 articles excluded:
- Not a full publication or full text not available: 78
- Not available in English: 8
- Not original data from an RCT, SR/MA, or observational study with comparator: 28
- Observational study of < 100 subjects: 48
- Not a study population of interest: 49
- No comparator of interest: 86
- No outcomes of interest: 189
- Outcomes not reported by underlying diagnosis: 771
- Does not meet study design criteria by outcome type: 491

1,909 passed abstract screening

161 articles representing 151 studies passed full-text screening and were included for abstraction

Data abstracted for 151 studies:\(^{a}\)
- KQ 1: 56 studies
- KQ 2: 7 studies
- KQ 3: 50 studies
- KQ 4: 8 studies
- KQ 5: 23 studies
- KQ 6: 5 studies
- Across All KQs: 21 studies

\(^{a}\)Some studies are relevant to more than one KQ.

Abbreviations: KQ=Key Question; RCT=randomized controlled trial; SR/MA=systematic review/meta-analysis
Description of Included Studies

Overall, we included 151 studies described in 161 publications: 56 studies were relevant to KQ 1, 7 studies to KQ 2, 50 studies to KQ 3, 8 studies to KQ 4, 23 studies to KQ 5, and 5 studies to KQ 6 (some studies were relevant to more than one KQ). Of the 151 included studies, 21 studies had adjusted their results for cause of infertility, but did not report their findings for specific causes of infertility and are discussed at the end of the results section. Globally the evidence supporting findings varying by patient characteristics such as age, ovarian reserve, race, BMI, and presence of other potential causes was minimal. We highlight in the report those cases where findings in these specific subgroups was possible.

Studies were conducted wholly or partly in continental Europe or the United Kingdom (52 studies, 34%), the United States or Canada (34 studies, 23%), the Middle East (32 studies, 21%), Asia (19 studies, 12%), Africa (10 studies, 7%), and other locations (Latin America [1 study; this study also had sites in the UK/Europe] and Australia/New Zealand [3 studies], total 2%). Appendix C provides a detailed listing of included articles. Appendix D provides a complete list of articles excluded at the full-text screening stage, with reasons for exclusion. Further details on the studies included for each KQ are provided in the relevant results sections, below, and in Appendix E. Detailed risk of bias information for each included study is reported in Appendix G.

Key Question 1. PCOS

KQ 1. What are the comparative safety and effectiveness of available treatment strategies for women with polycystic ovary syndrome (PCOS) who are infertile and who wish to become pregnant?

KQ 1a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, body mass index (BMI), presence of other potential causes of female infertility, or presence of male factor infertility?

Description of Studies Included for KQ 1 (PCOS)

We identified 61 articles describing 56 studies that addressed the comparative safety and effectiveness of available treatment strategies for infertility in women with PCOS. Four studies were described in nine publications, as follows:

- PCOS Study: Legro, 2007: Primary report and a companion paper
- PCOS 2 Study: Legro, 2014: Primary report and a companion paper
- Mutsaerts, 2016: Primary report and two companion papers
- Nahuis, 2011: Primary report and a companion paper

Of the 56 included studies, 52 were randomized controlled trials (RCTs). Twenty five of these were good quality, 25 were fair quality, and 2 were poor quality. In addition, we identified three good-quality observational studies and one fair-quality observational study.

Much of the research was done abroad, in subspecialty clinics and funded by an unclear or unknown sources. The breakdown of geographic location was: 5 studies in Africa, 11 studies in Asia, 1 study in Australia/NZ, 17 studies in the United States or Canada, 34 studies, 21% of the studies were conducted wholly or partly in continental Europe or the United Kingdom.
Middle East, 143,145,148,157,158,161,170,176,178,182,183,189-194 5 studies in the United States, 128,131,159,172,185 16 studies in the United Kingdom or continental Europe, 119,141,144,151,153,155,163-165,177,179-181,184,186,187 and 1 study in the UK/Europe and Latin America. 152 Settings included 2 studies conducted in general gynecology practices, 173,182 3 studies conducted in a hospital, 180,183,191 2 studies conducted in a combination of gynecological or subspecialty practices, 153,193 6 studies where the setting was unclear or not reported, 131,141,144,147,177,184 while the remaining 42 studies were conducted in subspecialty practices. Last, 8 studies reported government funding, 119,131,141,149,164,171,172,187 3 studies reported industry funding, 152,159,186 7 studies reported non-government, non-industry funding, 148,157,165,178,181,183,184 6 studies reported a combination of funding from a variety of sources, 128,146,160,174,179,188 while the remaining 33 studies did not report a funding source or it was unclear.

Further details on the characteristics of studies included for this KQ are provided in the following sections and Appendix E.

In addition to the above studies, seven systematic reviews; six good quality, 66,81,90,195-197 one fair quality, 78 addressed the comparative effectiveness of various treatments for infertility in women with PCOS are also discussed below and the consistency of their findings with our included studies are incorporated in to our strength of evidence ratings. In general, the randomized trials used standardized diagnostic criteria for PCOS, while the nonrandomized observational studies may have included other ovulatory disorders.

**Key Points for PCOS**

Key findings for outcomes in couples where the primary cause of infertility is PCOS include:

- Letrozole has a higher live birth rate than clomiphene citrate alone and lower multiple births, with no difference in ectopic pregnancy, or miscarriage (moderate for all outcomes), low birthweight, or time to pregnancy (low strength of evidence [SOE] for both these outcomes).
- Clomiphene citrate does not result in higher live birth rates compared with metformin (moderate SOE). Differences are also not found in the rates of multiple birth, ectopic pregnancy, or time to pregnancy (low SOE for all outcomes). There is a higher rate of miscarriage with combination clomiphene and metformin than clomiphene alone (low SOE)
- Letrozole or letrozole and berberine have a higher live birth rate than berberine alone (low SOE) with no difference in multiple births, miscarriage, or low birthweight rates (low SOE)
- There was no difference between clomiphene and tamoxifen for the outcomes of live birth or miscarriage (low SOE)
- There was no difference between laparoscopic ovarian drilling (LOD) and oral agents for live birth (moderate SOE) or miscarriage rates (low SOE). Multiple births were reduced given LOD (moderate SOE).
- Live birth (low SOE) and miscarriage rates (moderate SOE) did not differ between in vitro fertilization (IVF) treatment strategies.
- There was no difference in live birth rates for women who underwent lifestyle modification in combination with IVF compared with IVF alone (moderate SOE)
- There was no difference between type 1 diabetes mellitus diagnoses in children conceived with assisted reproductive technology (ART) compared to children conceived with no fertility treatment (moderate SOE)
As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE).

**Detailed Synthesis by Treatment for PCOS**

Included studies and their findings for the following treatments for PCOS are detailed in this section:

1. Oral Agents Alone
2. Oral Agents Alone Versus Acupuncture
3. Oral Agents Alone Versus Oral Agents With Intrauterine Insemination (IUI)
4. Oral Agents Alone Versus Surgical Management
5. Oral Agents Versus Gonadotropins
6. Lifestyle Interventions
7. Surgical Management Versus Gonadotropins
8. Gonadotropins With IUI
9. Assisted Reproductive Technology (ART)
   a. IVF
   b. Adjuncts to IVF
   c. Fresh Versus Frozen in IVF
   d. Intra-cytoplasmic sperm injection (ICSI)

**1. Oral Agents Alone for PCOS**

Oral agents used to induce ovulation in women with PCOS include selective estrogen receptor modulators (e.g., clomiphene citrate, tamoxifen), aromatase inhibitors (e.g., letrozole), dopamine agonists (e.g., cabergoline) and herbal medicines (e.g., berberine). We identified 12 randomized controlled trials (RCTs) (6 good quality, 6 fair quality) that addressed outcomes between different oral agents for ovulation induction without intrauterine insemination (IUI). These studies varied in the medication type used for oral ovulation as well as in adjunct treatments. These 12 studies included 3217 patients combined.

We also identified two systematic reviews that explored metformin versus clomiphene, one that focused on letrozole versus clomiphene, and one that examined clomiphene and other elective estrogen receptor modulators. Findings for our specific outcomes of interest are summarized in Table 4.

**Letrozole Versus Clomiphene**

Letrozole resulted in significantly more live births than clomiphene (moderate SOE). This finding was based on a good-quality meta-analysis by Franik et al. comparing clomiphene to letrozole, and which included the RCT by Legro et al. This meta-analysis of 9 RCTs (1783 subjects) reported a significant increase in live birth rate with letrozole (pooled odds ratio [OR] 1.64; 95% confidence interval [CI], 1.32 to 2.04). One additional recent good-quality RCT by Amer et al. reported live birth rates that were non-significantly higher in the letrozole group compared to the clomiphene group (48.8% versus 35.4%, p=0.089).

Multiple births was explored in the meta-analysis demonstrating a lower rate of multiple pregnancies with letrozole compared to clomiphene (OR 0.38, 95% CI, 0.17 to 0.84) based on 11 studies with 2385 subjects (moderate SOE). Our 3 included RCTs showed no difference in multiple births with very low rates of the outcome across the studies.
The meta-analysis did not report on ectopic pregnancies. The three individual RCTs reported no significant difference in ectopic pregnancy between the clomiphene and letrozole arms and a meta analysis of these three studies representing 886 women showed no difference (OR = 0.72, 95% CI 0.24 to 2.17) (SOE moderate).131,170,179

Miscarriage was reported in all 3 studies and in the meta-analysis by Franik and colleagues, with several definitions of miscarriage utilized. From the meta analysis, 2385 patients from 12 trials demonstrated no difference between letrozole and clomiphene in miscarriage rates (pooled OR 1.32, 95% CI 0.92 to 1.88) (moderate SOE).66 The two additional individual RCTs not included in the meta analysis comprised an additional 136 women and also reported no significant differences in miscarriage between the treatment arms.170,179

Neonatal death was reported in one good-quality RCT,131 which compared clomiphene to letrozole. There was not a significant difference between clomiphene and letrozole for cases of neonatal death although given the rarity of the outcome, much larger data sets are needed (insufficient SOE).

Finally, one good-quality RCT reported time to pregnancy for clomiphene versus letrozole.131 It did not show a significant difference in mean days to pregnancy between study arms (low SOE).

**Metformin Versus Clomiphene**

Five studies compared clomiphene to metformin (or metformin in combination with clomiphene).128,151,166,173 Three evaluated live birth as an outcome. One RCT151 demonstrated a significantly greater live birth rate with the combination of metformin and clomiphene compared to clomiphene alone, while another RCT173 reported non-significantly higher live birth rates with the combination of metformin and clomiphene than either metformin alone or clomiphene alone. The third RCT128 resulted in greater live births with clomiphene compared to metformin but no difference between clomiphene and combination therapy (Table 4).

Two meta-analyses,78,199 one good quality,199 and one fair quality,78 also compared metformin and clomiphene. Both meta-analyses reported no significant difference in live birth rates between women treated with metformin and clomiphene. The fair-quality meta-analysis by Sun et al.,78 reported an OR of 0.89; 95% CI, 0.71 to 1.13 in 4 RCTs (1012 patients) and the good-quality meta-analysis199 reported an OR of 0.71, 95% CI, 0.49 to 1.01 based on five studies. The latter meta-analysis reported evidence of heterogeneity by BMI, with live birth rates lower in the metformin group among obese women (OR 0.30, 95% CI, 0.17 to 0.52) but higher in the non-obese group (OR 1.71, 95% CI, 1.00 to 2.94).

The 2017 meta-analysis included four of our identified RCTs.128,151,166,173 Both meta-analyses also reported no difference in live birth rates comparing metformin plus clomiphene to clomiphene alone.78,199 Sun reported an OR of 0.99; 95% CI, 0.84 to 1.17) based on 3 studies with 912 patients78 and Morley et al. reported an OR of 1.21, 95% CI, 0.92 to 1.59 based on 9 studies with 1079 women.199 Together we rated the SOE as moderate for no difference in live birth rates.

Miscarriages were reported as an outcome in four of the RCTs128,160,166,173 and synthesized in the 2017 meta-analysis.199 The meta-analysis (which included three128,166,173 of the four individual included RCTs) reported no differences in miscarriage rates between metformin and clomiphene, whereas one RCT reported reduced miscarriages in the clomiphene group.160 The meta-analysis reported an increased risk for miscarriage in women taking clomiphene plus metformin compared to clomiphene alone (OR 1.59, 95% CI, 1.03 to 2.46) based on 9 studies
with 1096 women. None of the individual studies reported a statistically significant difference between the groups, although in each of the studies the miscarriage rate was higher in the clomiphene plus metformin group than the clomiphene group. Together these studies supported a low SOE of a higher rate of miscarriage in the combined therapy group.

The 2017 meta-analysis and our included studies did not support a difference in multiple births between clomiphene and metformin. Given the imprecision in these findings and suspected reporting bias of the included studies the strength of evidence was rated as low. Our included studies also did not support a difference in ectopic pregnancy (low SOE for both outcomes).

For the outcomes of congenital anomalies, the anomalies reported in these studies varied in severity and type and given the rarity of outcomes and imprecise evidence the SOE was rated as insufficient. None of the studies however found significant differences between intervention groups.

One good-quality RCT reported time to pregnancy for clomiphene versus metformin or placebo in combination with clomiphene. It did not show a significant difference in mean days to pregnancy between study arms (low SOE).

Tamoxifen Versus Clomiphene

One fair-quality RCT compared tamoxifen with clomiphene. Live births and miscarriages were not significantly different between the treatment arms. A meta-analysis also reported on tamoxifen versus clomiphene and found no significant differences in live births (OR 1.24, 95% CI 0.59-2.62) in two studies with a total of 195 women or in miscarriages (OR 1.81, 95% CI 0.80-4.12) in four studies with a total of 653 women. Within that meta analysis, both outcomes were judged to have a low grade of evidence. We rated the strength of evidence for both outcomes as low for no difference between tamoxifen and clomiphene.

Cabergoline Versus Clomiphene

One fair-quality RCT compared clomiphene alone to clomiphene plus cabergoline. Live birth rates were not reported, and no significant differences in miscarriages or multiple pregnancies were reported. SOE was rated as insufficient given findings from one small study with potential limitations.

Letrozole Versus Berberine

One good-quality RCT compared letrozole, berberine and a combination of the two. Live births were similar between the letrozole and combination arms (36.3% and 34.4%, p=0.69), and both of these arms had significantly higher live birth rates than the berberine arm (22.0%, p=0.001) (low SOE). No significant differences between treatment arms were reported for multiple births, miscarriage or birthweight (low SOE for all three outcomes).

There was no evidence regarding costs, short-term adverse effects, and long-term child or maternal outcomes.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Legro, 2014(^{131}) RCT (750)</td>
<td>CC</td>
<td>Letrozole</td>
<td>72/376 (19.1) (15.3 to 23.3)</td>
<td>103/374 (27.5) (23.1 to 32.2)</td>
<td>0.007</td>
<td>Greater live births with letrozole compared to clomiphene</td>
</tr>
<tr>
<td></td>
<td>Amer, 2017(^{139}) RCT (159)</td>
<td>CC</td>
<td>Letrozole</td>
<td>28/79 (35.4)</td>
<td>39/80 (48.8)</td>
<td>0.089</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Morin-Papunen, 2012(^{81}) RCT (320)</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>46/160 (28.9) (22.0 to 36.0)</td>
<td>66/160 (41.1) (33.8 to 38.9)</td>
<td>0.03</td>
<td>Greater live births with combination of metformin and clomiphene compared to clomiphene alone</td>
</tr>
<tr>
<td></td>
<td>Legro, 2007(^{28}) RCT (626)</td>
<td>CC</td>
<td>Metformin</td>
<td>47/209 (22.5) (17.1 to 28.4)</td>
<td>15/208 (7.2) (4.1 to 11.1)</td>
<td>&lt;0.001</td>
<td>Greater live births with clomiphene compared to metformin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metformin+ CC</td>
<td>47/209 (22.5) (17.1 to 28.4)</td>
<td>56/209 (26.8) (21.0 to 33.0)</td>
<td>0.31</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kar, 2015(^{173}) RCT (105)</td>
<td>CC</td>
<td>Metformin</td>
<td>9/32 (28.1)</td>
<td>9/24 (37.5)</td>
<td>0.46</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Metformin+ CC</td>
<td>9/32 (28.1)</td>
<td>10/24 (41.6)</td>
<td>0.29</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Topçu, 2017(^{191}) RCT (88)</td>
<td>CC</td>
<td>Tamoxifen</td>
<td>6/46 (13.0)</td>
<td>8/42 (19.0)</td>
<td>0.617</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Wu, 2016(^{175}) RCT (644)</td>
<td>Letrozole</td>
<td>Berberine</td>
<td>78/215 (36.3)</td>
<td>47/214 (22.0)</td>
<td>0.001</td>
<td>Greater live births for letrozole or letrozole + berberine compared with berberine alone. No difference between letrozole and letrozole + berberine</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Berberine + Letrozole</td>
<td>78/215 (36.3)</td>
<td>74/215 (34.4)</td>
<td>0.687</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Study Design (N Patients)</td>
<td>Interventions</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>---------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>----------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Legro, 2014</td>
<td>CC Letrozole</td>
<td>5/376 (1.3) (0.4 to 2.7)</td>
<td>4/374 (1.1) (0.3 to 2.3)</td>
<td>0.175</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Ghahiri, 2016</td>
<td>CC Letrozole</td>
<td>0/24 (0)</td>
<td>0/24 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amer, 2017</td>
<td>CC Letrozole</td>
<td>0/34 (0.0)</td>
<td>3/49 (6.1)</td>
<td>0.201</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson, 2010</td>
<td>CC Metformin</td>
<td>1/36 (2.8) (0.1 to 10.0)</td>
<td>1/35 (2.9) (0.1 to 10.3)</td>
<td>0.98</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC Metformin+</td>
<td>1/36 (2.8) (0.1 to 10.0)</td>
<td>1/35 (2.9) (0.1 to 10.3)</td>
<td>0.98</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC Letrozole</td>
<td>0/34 (0.0)</td>
<td>3/49 (6.1)</td>
<td>0.201</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009</td>
<td>CC Metformin</td>
<td>0/39 (0)</td>
<td>0/38 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC Metformin+</td>
<td>0/39 (0)</td>
<td>0/38 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2007</td>
<td>CC Metformin</td>
<td>3/50 (6.0) (1.3 to 14.0)</td>
<td>0/18 (0)</td>
<td>0.29</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC Metformin+</td>
<td>3/50 (6.0) (1.3 to 14.0)</td>
<td>3/65 (4.6) (1.0 to 10.8)</td>
<td>0.74</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2018</td>
<td>CC CC + Cabergoline</td>
<td>1/8 (12.5)</td>
<td>2/19 (10.5)</td>
<td>0.83</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu, 2016</td>
<td>Letrozole Berberine</td>
<td>1/78 (1.2)</td>
<td>0/47 (0.0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Letrozole Berberine + Letrozole</td>
<td>1/78 (1.2)</td>
<td>3/74 (4.1)</td>
<td>0.357</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Legro, 2014</td>
<td>CC Letrozole</td>
<td>3/376 (0.8) (0.2 to 1.9)</td>
<td>4/374 (1.1) (0.3 to 2.3)</td>
<td>0.67</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Ghahiri, 2016</td>
<td>CC Letrozole</td>
<td>2/24 (8.3)</td>
<td>3/29 (10.3)</td>
<td>0.80</td>
<td>No difference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention ( N ) (%)(95% CI)</td>
<td>Results Comparator ( N ) (%)(95% CI)</td>
<td>( P ) Value</td>
<td>Summary of Study Findings</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>--------------------------------</td>
<td>--------------------------------</td>
<td>----------------</td>
<td>--------------------------</td>
<td></td>
</tr>
<tr>
<td>Amer, 2017(^{179})</td>
<td>CC</td>
<td>Letrozole</td>
<td>0/34 (0.0)</td>
<td>1/49 (2.0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (159)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009(^{166})</td>
<td>CC</td>
<td>Metformin</td>
<td>0/39 (0)</td>
<td>0/38 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (124)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009(^{166})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>0/39 (0)</td>
<td>0/38 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Legro, 2007(^{128})</td>
<td>CC</td>
<td>Metformin</td>
<td>2/209 (1.0)</td>
<td>0/208 (0)</td>
<td>0.16</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (626)</td>
<td></td>
<td></td>
<td>(0.1 to 2.6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2007(^{128})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>2/209 (1.0)</td>
<td>2/209 (1.0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (626)</td>
<td></td>
<td></td>
<td>(0.1 to 2.6)</td>
<td>(0.1 to 2.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kar, 2015(^{173})</td>
<td>CC</td>
<td>Metformin</td>
<td>1/10 (10)</td>
<td>0/13 (0.0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (105)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kar, 2015(^{173})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>1/10 (10)</td>
<td>0/12 (0.0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Amer, 2017(^{179})</td>
<td>CC</td>
<td>Letrozole</td>
<td>30/103 (29.1)</td>
<td>49/154 (31.8)</td>
<td>0.65</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (750)</td>
<td></td>
<td></td>
<td>(20.8 to 38.2)</td>
<td>(24.7 to 39.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ghahiri, 2016(^{170})</td>
<td>CC</td>
<td>Letrozole</td>
<td>6/24 (25.0)</td>
<td>5/29 (17.2)</td>
<td>0.38</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amer, 2017(^{179})</td>
<td>CC</td>
<td>Letrozole</td>
<td>6/34 (17.6)</td>
<td>9/49 (18.4)</td>
<td>0.93</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (159)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2007(^{128})</td>
<td>CC</td>
<td>Metformin</td>
<td>16/62 (25.8)</td>
<td>10/25 (40.0)</td>
<td>0.19</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (626)</td>
<td></td>
<td></td>
<td>(15.8 to 37.3)</td>
<td>(22.1 to 59.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2007(^{128})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>16/62 (25.8)</td>
<td>24/80 (30.0)</td>
<td>0.58</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (626)</td>
<td></td>
<td></td>
<td>(15.8 to 37.3)</td>
<td>(20.5 to 40.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009(^{166})</td>
<td>CC</td>
<td>Metformin</td>
<td>0/39 (0)</td>
<td>0/38 (0)</td>
<td>NS</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>RCT (124)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009(^{166})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>0/39 (0)</td>
<td>1/38 (2.6)</td>
<td>0.31</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Johnson, 2010(^{160})</td>
<td>CC</td>
<td>Metformin</td>
<td>0/36 (0)</td>
<td>4/35 (11.4)</td>
<td>0.037</td>
<td>Reduced miscarriage with clomiphene compared to metformin</td>
<td></td>
</tr>
<tr>
<td>RCT (171)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johnson, 2010(^{160})</td>
<td>CC</td>
<td>Metformin+ CC</td>
<td>0/36 (0)</td>
<td>3/35 (8/6)</td>
<td>0.073</td>
<td>No difference</td>
<td></td>
</tr>
</tbody>
</table>

**Pregnancy complications: Miscarriage**

- Legro, 2014\(^{131}\) CC Letrozole 30/103 (29.1) (20.8 to 38.2) 49/154 (31.8) (24.7 to 39.4) 0.65 No difference
- Ghahiri, 2016\(^{170}\) CC Letrozole 6/24 (25.0) 5/29 (17.2) 0.38 No difference
- Amer, 2017\(^{179}\) CC Letrozole 6/34 (17.6) 9/49 (18.4) 0.93 No difference
- Legro, 2007\(^{128}\) CC Metformin 16/62 (25.8) (15.8 to 37.3) 10/25 (40.0) (22.1 to 59.4) 0.19 No difference
- Zain, 2009\(^{166}\) CC Metformin 0/39 (0) 0/38 (0) NS No difference
- Zain, 2009\(^{166}\) CC Metformin+ CC 0/39 (0) 1/38 (2.6) 0.31 No difference
- Johnson, 2010\(^{160}\) CC Metformin 0/36 (0) 4/35 (11.4) 0.037 Reduced miscarriage with clomiphene compared to metformin

Metformin+ CC 0/36 (0) 3/35 (8/6) 0.073 No difference
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neonatal outcomes: Birthweight (kg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>Letrozole</td>
<td></td>
<td>Mean = 3.23 (+/- 0.7153)</td>
<td>Mean = 3.23 (+/- 0.6574)</td>
<td>0.83</td>
<td>No difference</td>
</tr>
<tr>
<td>Wu, 2016&lt;sup&gt;175&lt;/sup&gt;</td>
<td>Letrozole</td>
<td>Berberine</td>
<td></td>
<td>Mean = 3.463 (+/- 0.575 SD)</td>
<td>Mean = 3.542 (+/- 0.399 SD)</td>
<td>0.216</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>RCT (644)</td>
<td>Berberine + Letrozole</td>
<td></td>
<td>Mean = 3.463 (+/- 0.575 SD)</td>
<td>Mean = 3.484 (+/- 0.504 SD)</td>
<td>0.246</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CC</td>
<td>Letrozole</td>
<td></td>
<td>1/66 (1.5) (0.0 to 5.5)</td>
<td>4/102 (3.9) (1.1 to 8.4)</td>
<td>0.37</td>
<td>No difference</td>
</tr>
<tr>
<td>Amer, 2017&lt;sup&gt;179&lt;/sup&gt;</td>
<td>CC</td>
<td>Letrozole</td>
<td></td>
<td>0</td>
<td>0</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Legro, 2007&lt;sup&gt;128&lt;/sup&gt;</td>
<td>CC</td>
<td>Metformin</td>
<td></td>
<td>0/209 (0)</td>
<td>0/208 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>RCT (626)</td>
<td>Metformin+ CC</td>
<td></td>
<td>0/209 (0)</td>
<td>2/209 (1.0) (0.1 to 2.6)</td>
<td>0.15</td>
<td>No difference</td>
</tr>
<tr>
<td>Time to pregnancy (days)</td>
<td>Legro, 2014&lt;sup&gt;131&lt;/sup&gt;</td>
<td>CC</td>
<td>Letrozole</td>
<td>85.9</td>
<td>90.4</td>
<td>0.27</td>
<td>No difference</td>
</tr>
</tbody>
</table>
### 2. Oral Agents Alone Versus Acupuncture for PCOS

One fair-quality factorial RCT with 1000 women\(^\text{188}\) compared clomiphene and acupuncture alone and combined. Results for live births, miscarriage, ectopic pregnancy and congenital abnormality are summarized in Table 5. When analyzing the main effects of acupuncture and clomiphene, the live birth rate was significantly higher in the clomiphene group as compared to placebo whereas it was not significantly different for the active and control acupuncture. SOE was rated as low for all outcomes given findings from one study with potential risk of bias.

#### Table 5. Outcomes for comparisons of oral agents alone versus acupuncture in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Wu, 2017(^\text{188}) RCT (1000)</td>
<td>Active Acupuncture + CC</td>
<td>Control Acupuncture + CC</td>
<td>69/235 (29.4)</td>
<td>66/236 (28.0)</td>
<td>0.73</td>
<td>Live birth rates significantly higher for clomiphene vs. placebo; not significantly different for active vs. control acupuncture</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic Pregnancy</td>
<td>Wu, 2017(^\text{188}) RCT (1000)</td>
<td>Active Acupuncture + CC</td>
<td>Control Acupuncture + CC</td>
<td>1/108 (0.9)</td>
<td>0/106 (1.9)</td>
<td>0.54</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Wu, 2017(^\text{188}) RCT (1000)</td>
<td>Active Acupuncture + CC</td>
<td>Control Acupuncture + CC</td>
<td>38/108 (35.2)</td>
<td>37/106 (34.9)</td>
<td>0.96</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes:</td>
<td>Wu, 2017(^\text{188})</td>
<td>Active Acupuncture + CC</td>
<td>Control Acupuncture + CC</td>
<td>1/69 (1.4)</td>
<td>0/66 (0.0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; N=number of patients; NS= not statistically significant; PCOS=polycystic ovary syndrome; RCT=randomized control trial; SD=standard deviation.
### 3. Oral Agents Alone Versus Oral Agents With IUI for PCOS

One good-quality RCT\(^\text{157}\) compared clomiphene citrate without IUI to clomiphene citrate with IUI in women with PCOS. The results for live birth (reported as any live birth per patient), pregnancy complications (multiple births, ectopic, miscarriage) were reported. Results for live births and pregnancy complications are summarized in Table 6 (insufficient SOE). There was no evidence for this treatment regarding neonatal outcomes, time to pregnancy, costs, short-term adverse effects, and long-term child or maternal outcomes.

A second good-quality RCT\(^\text{182}\) compared Myo-inositol before and during ovulation induction followed by IUI compared to ovulation induction and IUI alone in women with PCOS. Frequency of OHSS and multiple births were reported and are summarized in Table 6 (insufficient SOE).

#### Table 6. Outcomes for oral agents alone versus oral agents with IUI in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Live birth: Any/patient</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2011(^\text{157}) RCT (188)</td>
<td>CC + IUI</td>
<td>CC + timed Intercourse</td>
<td>18/93 (19.4%) (12.0% to 27.9%)</td>
<td>17/95 (17.9%) (10.9% to 26.2%)</td>
<td>0.33</td>
<td>No difference</td>
</tr>
<tr>
<td><strong>Pregnancy complications: Multiple births</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2011(^\text{157}) RCT (188)</td>
<td>CC + IUI</td>
<td>CC + timed Intercourse</td>
<td>2/22 (9.1%) (1.2% to 23.8%)</td>
<td>2/21 (9.5%) (1.2% to 24.9%)</td>
<td>0.46</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; N=number of patients; NS= not statistically significant; PCOS=polycystic ovary syndrome; RCT=randomized control trial; SD=standard deviation.
### Summary of Study Findings

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Emekci Ozay 2017&lt;sup&gt;182&lt;/sup&gt; RCT (196)</td>
<td>MYO + COH/IUI</td>
<td>COH/IUI</td>
<td>2/16 (12.5%)</td>
<td>1/11 (9.1%)</td>
<td>0.78</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Abu Hashim, 2011&lt;sup&gt;157&lt;/sup&gt; RCT (188)</td>
<td>CC + IUI</td>
<td>CC + timed Intercourse</td>
<td>0/93 (0%)</td>
<td>0/95 (0%)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS</td>
<td>Emekci Ozay 2017&lt;sup&gt;182&lt;/sup&gt; RCT (196)</td>
<td>MYO + IUI</td>
<td>IUI alone</td>
<td>1/86 (1.2%)</td>
<td>3/90 (3.3%)</td>
<td>0.35</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; COH=controlled ovarian hyperstimulation; IUI=intruterine insemination; MYO=Myo-inositol; N=number of patients; NR=not reported; NS= not statistically significant; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome; RCT=randomized control trial; SD=standard deviation

### 4. Oral Agents Alone Versus Surgical Management for PCOS

Seven RCTs<sup>154,156,163,165,168,169,189</sup> (four good quality, three fair quality) compared the use of oral agents without IUI to surgical management in women with PCOS. These studies varied in the medication used as well as in the specific surgical methods. These 7 studies included 918 patients. One good-quality systematic review of 25 studies also explored laparoscopic ovarian drilling versus oral agents alone.<sup>81</sup> All but one<sup>189</sup> of the seven included studies were included in this systematic review and therefore the findings of this systematic review are highlighted when they assessed outcomes of interest.

Outcomes are summarized in Table 7 and demonstrate that there were no significant differences in treatment outcomes between pharmacologic and surgical approaches for live birth within the individual studies. The systematic review of 25 studies by Farquhar et al.<sup>81</sup> included 8 studies (1034 women) where live birth was evaluated also found no significant differences in live birth rates comparing LOD to oral agents (OR = 0.77, 95% CI 0.59 to 1.01) (moderate SOE). When exploring specific oral agents, this review found that LOD compared to clomiphene plus tamoxifen (OR 0.81; 95% CI, 0.42 to 1.53 based on 1 trial, n=150), or to letrozole (OR 0.97; 95% CI, 0.59 to 1.59 based on 3 trials, n=318) did not support a difference in live birth rates, but it did find a significantly lower birth rate with LOD compared to clomiphene plus metformin (OR 0.44; 95% CI, 0.24 to 0.82 based on 2 trials, n=159).

Our included studies did not demonstrate a difference in miscarriage. The systematic review synthesized findings from 15 studies (1592 women) and found no difference between LOD and
other treatments (OR 1.1, 95% CI 0.74 to 1.61) (low SOE). The evidence was downgraded since there were inadequate explanations of randomization (in 3 trials), allocation concealment (8 trials) and inadequate or no blinding reported in 8 trials. The one study not included in the systematic review also showed no significant difference in miscarriage (18% vs. 7%, p=0.41). The multiple pregnancy rate was evaluated in 12 studies (1129 women) and found to be lower in women undergoing LOD (OR 0.21, 95% CI 0.08 to 0.58) (moderate SOE).

The evidence did not show any differences between treatments for the outcomes of OHSS, and surgical complications but was considered insufficient SOE given the small numbers of events and the imprecision in the included studies.

There was no evidence for this treatment regarding neonatal outcomes, time to pregnancy, costs, and long-term child or maternal outcomes.

Table 7. Outcomes for oral agents alone versus surgical management in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention (%) (95% CI)</th>
<th>Results Comparator (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Abu Hashim, 2011&lt;sup&gt;154&lt;/sup&gt; RCT (176)</td>
<td>CC</td>
<td>LOD</td>
<td>25/89 (28.1) (19.3 to 37.8)</td>
<td>28/87 (32.2) (22.8 to 42.3)</td>
<td>0.55</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Palomba, 2010&lt;sup&gt;163&lt;/sup&gt; RCT (47)</td>
<td>CC</td>
<td>LOD</td>
<td>12/23 (52.2) (32.2 to 71.8)</td>
<td>13/24 (54.2) (34.5 to 73.2)</td>
<td>1</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Amer, 2009&lt;sup&gt;165&lt;/sup&gt; RCT (65)</td>
<td>CC</td>
<td>LOD</td>
<td>18/32 (56.3) (39.1 to 72.7)</td>
<td>15/33 (45.5) (29.1 to 62.3)</td>
<td>0.27</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2010&lt;sup&gt;168&lt;/sup&gt; RCT (260)</td>
<td>Letrozole</td>
<td>LOD</td>
<td>32/128 (25.0) (17.9 to 32.8)</td>
<td>33/132 (25.0) (18.0 to 32.7)</td>
<td>1</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Zakherah, 2010&lt;sup&gt;169&lt;/sup&gt; RCT (150)</td>
<td>CC + tamoxifen</td>
<td>LOD</td>
<td>37/75 (49.3) (38.1 to 60.6)</td>
<td>33/75 (44.0) (33.0 to 55.3)</td>
<td>0.35</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Abbisallah, 2011&lt;sup&gt;156&lt;/sup&gt; RCT (140)</td>
<td>Letrozole</td>
<td>LOD</td>
<td>0/23 (0)</td>
<td>0/16 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-----------------------------------</td>
<td>----------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2010 [168]</td>
<td>Letrozole</td>
<td>LOD</td>
<td>0/128 (0)</td>
<td>0/132 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Amer, 2009 [163]</td>
<td>CC</td>
<td>LOD</td>
<td>0/20 (0)</td>
<td>0/17 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2011 [154]</td>
<td>CC</td>
<td>LOD</td>
<td>5/30 (16.7) (5.8 to 31.7)</td>
<td>6/34 (17.6) (7.0 to 31.6)</td>
<td>0.92</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Abdellah, 2011 [156]</td>
<td>Letrozole</td>
<td>LOD</td>
<td>2/25 (8.0) (1.0 to 2.1)</td>
<td>4/20 (20.0) (6.1 to 39.6)</td>
<td>0.231</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Amer, 2009 [165]</td>
<td>CC</td>
<td>LOD</td>
<td>2/20 (10.0) (1.3 to 26.0)</td>
<td>2/17 (11.8) (1.6 to 30.2)</td>
<td>0.62</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Abu Hashim, 2010 [168]</td>
<td>Letrozole</td>
<td>LOD</td>
<td>4/128 (3.1) (0.9 to 6.7)</td>
<td>4/132 (3.0) (0.8 to 6.5)</td>
<td>0.92</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Zakherah, 2010 [169]</td>
<td>CC + tamoxifen</td>
<td>LOD</td>
<td>3/40 (7.5) (1.6 to 17.3)</td>
<td>5/38 (13.2) (4.5 to 25.4)</td>
<td>0.06</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Ibrahim, 2017 [169]</td>
<td>Letrozole</td>
<td>LOD</td>
<td>1/14 (7)</td>
<td>2/11 (18)</td>
<td>0.41</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Amer, 2009 [165]</td>
<td>CC</td>
<td>LOD</td>
<td>1/32 (3.1) (0.1 to 11.2)</td>
<td>0/33 (0)</td>
<td>0.31</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Pregnancy complications: Miscarriage

Short-term adverse effects: OHSS
5. Oral Agents Versus Gonadotropins for PCOS

Five RCTs, three good-quality and two fair-quality, compared clomiphene to follicle-stimulating hormone (FSH) in women with PCOS with variation across the studies in the treatment protocols. Treatment with clomiphene citrate was compared to FSH only in two studies. Another study with a factorial design compared FSH with or without IUI to CC with or without IUI. The fourth study compared clomiphene plus urinary FSH to FSH alone and a fifth study compared extended clomiphene treatment to human menopausal gonadotropin. Outcomes for the studies are summarized in Table 8.

Higher birth rates and higher miscarriage rates were reported with FSH treatment compared to CC in one good-quality RCT, whereas no differences between FSH and CC were observed in two other studies, one good-quality and one fair quality. SOE was rated as insufficient for both outcomes given inconsistent evidence from the included studies. Ectopic pregnancies and birthweight did not significantly differ between groups (SOE low).

A good-quality RCT compared clomiphene plus urinary FSH to FSH alone. No statistically significant difference in live birth or pregnancy complications were reported. Given the limited evidence and imprecise findings these outcomes were rated insufficient for strength of evidence.

There was no evidence for these treatment comparisons regarding time to pregnancy, costs, short-term adverse effects, and long-term child outcomes.

One RCT compared letrozole to FSH among women with clomiphene resistant PCOS. Outcomes of interested that were reported include miscarriage rate and OHSS development with no difference seen between groups.

We also identified one meta-analysis of five studies with 264 women that compared ovulation induction using gonadotropins with and without metformin. Live births were significantly higher in the metformin group 457/1000 versus 267/1000 (OR 2.31, 95% CI 1.24 to 4.33) based on two studies including 180 women. Miscarriage rates were not significantly
different, based on three studies including 84 women (OR 0.62, 95% CI 0.19 to 2.01). OHSS also was not significantly different based on two studies including 180 women (OR 0.32, 95% CI 0.01 to 8.23). The authors of the meta-analysis judged the overall SOE to be low.

Table 8. Outcomes for oral agents versus gonadotropins in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Homburg, 2012 (302)</td>
<td>CC</td>
<td>FSH</td>
<td>53/143 (37.1) (29.4 to 45.1)</td>
<td>72/159 (45.3) (37.6 to 53.0)</td>
<td>0.12</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Hossein-Rashidi, 2016 (104)</td>
<td>CC</td>
<td>FSH</td>
<td>5/52 (9.6)</td>
<td>5/44 (11.4)</td>
<td>0.78</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Weiss, 2018 (666)</td>
<td>CC</td>
<td>CC+IUI</td>
<td>66/171 (38.6)</td>
<td>72/163 (44.2)</td>
<td>FSH vs. CC: 0.0124</td>
<td>Greater live births for FSH than CC; no significant difference between IUI and intercourse.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FSH</td>
<td>78/163 (47.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FSH+IUI</td>
<td>89/164 (54.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ghanem, 2013 (159)</td>
<td>CC + FSH</td>
<td>FSH</td>
<td>22/87 (25.3) (16.8 to 34.9)</td>
<td>19/87 (21.8) (13.9 to 31.0)</td>
<td>0.85</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Homburg, 2012 (302)</td>
<td>CC</td>
<td>FSH</td>
<td>0/143 (0)</td>
<td>2/159 (1.3) (0.2 to 3.5)</td>
<td>0.17</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Homburg, 2012 (302)</td>
<td>CC</td>
<td>FSH</td>
<td>1/143 (0.7) (0.0 to 2.6)</td>
<td>1/159 (0.6) (0.0 to 2.3)</td>
<td>0.91</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Weiss, 2018 (666)</td>
<td>CC</td>
<td>CC+IUI</td>
<td>1/171 (0.6)</td>
<td>3/163 (1.8)</td>
<td>0.31</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FSH</td>
<td>1/163 (0.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>FSH+IUI</td>
<td>1/164 (0.6)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Homburg, 2012 (302)</td>
<td>CC</td>
<td>FSH</td>
<td>5/143 (3.5) (1.2 to 7.1)</td>
<td>7/159 (4.4) (1.8 to 8.1)</td>
<td>0.68</td>
<td>No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>----------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>FSH</td>
<td>0/52 (0.0)</td>
<td>1/44 (2.3)</td>
<td>0.27</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CC</td>
<td>CC+IUI</td>
<td>3/171 (1.8)</td>
<td>8/163 (4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FSH</td>
<td></td>
<td>9/163 (5.5)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FSH+IUI</td>
<td></td>
<td>15/164 (9.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Letrozole</td>
<td>FSH</td>
<td>2/21</td>
<td>3/24</td>
<td>0.999</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Extended CC</td>
<td>FSH</td>
<td>5/160</td>
<td>4/158</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight (kg)</td>
<td></td>
<td>CC</td>
<td>CC+IUI</td>
<td>3.408 (0.491 SD)</td>
<td>3.178 (0.714 SD)</td>
<td></td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FSH</td>
<td></td>
<td>3.302 (0.769 SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FSH+IUI</td>
<td></td>
<td>3.279 (0.695 SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; FSH=follicle-stimulating hormone; IUI=intrauterine insemination; N=number of patients; NR=not reported; NS=not statistically significant; PCOS=polycystic ovary syndrome; RCT=randomized control trial; SD=standard deviation
6. Lifestyle Interventions for PCOS

Three studies looked at methods of lifestyle modifications for women with PCOS and their impact on outcomes of interest. One good-quality 3-arm study\textsuperscript{172} compared preconception continuous oral contraceptives (OCPs) to lifestyle modification with caloric restriction, weight loss medication and increased physical activity to a combined treatment of OCPs and lifestyle modifications. After the preconception intervention, all women started standard ovulation induction for four cycles with clomiphene citrate. The primary outcome was live birth rate, and relevant secondary outcomes included fecundity per ovulated patient, and adverse outcomes (ectopic pregnancy). Results are summarized in Table 9. There was no evidence for difference in live birth rate between arms, though fecundity per patient who ovulated was higher in women randomized to lifestyle intervention compared to OCP alone.

One good-quality multicenter RCT (LIFEstyle study)\textsuperscript{119} conducted a predetermined subgroup analyses based on ovulatory status (anovulatory vs. ovulatory). The intervention consisted of a 6-month program aimed at loss of 5-10\% of original body weight. Of those women who were anovulatory, 76\% of women in the intervention group (pre-treatment lifestyle) and 74\% in the control group (prompt treatment) met criteria for PCOS. Outcomes reported include overall and healthy live birth <24 months. There were no significant differences between lifestyle intervention and control on healthy live birth rate or overall live birth rate between ovulatory and anovulatory women. In addition, the effect of lifestyle intervention on overall and healthy birth rate was not altered by ovulatory status.

Finally, a third good-quality multi-center RCT\textsuperscript{181} compared a pre-IVF treatment 12 week strict low calorie liquid formula diet (LCD) to IVF treatment alone. 23\% of intervention participants and 18.3\% of IVF only participants had PCOS, and a subgroup analysis was conducted for this group of 81 women (Table 9). No difference was seen between arms in the live birth rates. Combined, these three studies supported a moderate SOE for no difference in live birth rates from lifestyle modification interventions.

Table 9. Outcomes for lifestyle intervention versus oral contraceptive pills or no intervention in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Legro 2015\textsuperscript{172} RCT (132)</td>
<td>Lifestyle + CC</td>
<td>OCP + CC</td>
<td>13/50 (26.0)</td>
<td>5/49 (10.2)</td>
<td>0.06</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OCP + Lifestyle + CC</td>
<td></td>
<td>12/50 (24.0)</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mutsaerts, 2016 \textsuperscript{119} RCT (564)</td>
<td>Lifestyle</td>
<td>No Lifestyle</td>
<td>72/123 (58.5)</td>
<td>83/140 (59.3)</td>
<td>0.90</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Einarsson, 2017\textsuperscript{181} RCT (317)</td>
<td>LCD +IVF</td>
<td>IVF alone</td>
<td>11/40 (27.5)</td>
<td>9/41 (22.0)</td>
<td>0.75</td>
<td>No difference</td>
</tr>
</tbody>
</table>
### 7. Surgical Management Versus Gonadotropins for PCOS

Three articles reporting two fair-quality RCTs\(^1\text{41,142,148}\) compared surgical management to gonadotropins with IUI for PCOS. These studies varied in the medication type/protocol used for oral ovulation induction as well as in the specific methods used for laparoscopic ovarian electrocauterization (a form of laparoscopic ovarian drilling) and the protocol following surgery. The two studies included 272 patients.

Results are summarized in Table 10. In the study reporting live birth,\(^\text{141}\) rates were not significantly different between treatments. Multiple births and costs were also not different between strategies.\(^\text{142}\) In a study comparing adverse outcomes, OHSS was reported in 3.8% of subjects receiving hMG, and no cases in the laparoscopic electrocauterization arm.\(^\text{148}\) Given the evidence from one small fair-quality study for each of these outcomes it was graded as insufficient. There was no evidence for this treatment regarding neonatal outcomes, time to pregnancy, and long-term child or maternal outcomes.

#### Table 10. Outcomes for surgical management versus gonadotropins in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Nahuis, 2011(^\text{141}) RCT (168)</td>
<td>LEC</td>
<td>rFSH</td>
<td>71/83 (85.5) (77.3 to 92.2)</td>
<td>69/85 (81.2) (72.3 to 88.7)</td>
<td>0.63</td>
<td>No difference</td>
</tr>
<tr>
<td>Live birth: Per-ovulated patient</td>
<td>Legro 2015(^\text{142}) RCT (132)</td>
<td>Lifestyle + CC</td>
<td>OCP + CC</td>
<td>13/36 (36.1)</td>
<td>5/36 (13.9)</td>
<td>0.04</td>
<td>Increased live birth rate for Lifestyle intervention than OCPs alone.</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Legro 2015(^\text{142}) RCT (132)</td>
<td>Lifestyle + CC</td>
<td>OCP + CC</td>
<td>1/16 (6.3)</td>
<td>0/8</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>OCP + Lifestyle + CC</td>
<td>0/14 (0)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CC = clomiphene citrate; CI = confidence interval; hMG = human menopausal gonadotropin; LCD = low calorie liquid formula diet; N = number of patients; NR = not reported; NS = not statistically significant; OCP = oral contraceptives; OHSS = ovarian hyperstimulation syndrome; PCOS = polycystic ovary syndrome; RCT = randomized control trial; rFSH = recombinant follicle-stimulating hormone.
### 8. Gonadotropins With IUI for PCOS

Two fair-quality RCTs\textsuperscript{143,159} examined the effects of gonadotropins with IUI. These studies varied in the medication type used for oral ovulation as well as in adjunct treatments. These two studies included 374 patients. Table 11 summarizes the findings for live birth, pregnancy complications, and short-term adverse effects. Live birth was investigated by both studies but used varying outcomes for measuring live birth. The strength of evidence was rated as insufficient given evidence from one fair-quality trial for either live birth measure. These same two studies examined pregnancy complications, reporting on miscarriage and multiple births. Neither study found significant differences between their intervention groups for either outcome. The imprecision of the findings and the quality of the included studies resulted in insufficient strength of evidence.

Finally, one fair-quality RCTs\textsuperscript{143} reported on short-term adverse effects of treatment, specifically OHSS (insufficient SOE). There was no evidence regarding neonatal outcomes, time to pregnancy, costs, and long-term child or maternal outcomes.

#### Table 11. Outcomes for gonadotropins with IUI in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/cycle</td>
<td>Stadtmauer, 2011\textsuperscript{139}</td>
<td>rFSH</td>
<td>rFSH + GnRH antagonist (flexible)</td>
<td>10/53 (18.9)</td>
<td>18/54 (33.3)</td>
<td>0.09</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>RCT (98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>rFSH</td>
<td>rFSH + GnRH antagonist (start day)</td>
<td>10/53 (18.9)</td>
<td>7/47 (12.3)</td>
<td>0.37</td>
<td>No difference</td>
</tr>
<tr>
<td>Live birth: Single/cycle</td>
<td>Rashidi, 2015\textsuperscript{143}</td>
<td>rFSH</td>
<td>hMG</td>
<td>21/132 (15.9)</td>
<td>14/144 (9.7)</td>
<td>0.14</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>RCT (276)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Summary of Study Findings

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Rashidi, 2015&lt;sup&gt;143&lt;/sup&gt; RCT (276)</td>
<td>rFSH</td>
<td>hMG</td>
<td>1/132 (0.8%) (0.0 to 2.8)</td>
<td>2/144 (1.4) (0.2 to 3.8)</td>
<td>0.62</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Stadtmauer, 2011&lt;sup&gt;159&lt;/sup&gt; RCT (98)</td>
<td>rFSH</td>
<td>rFSH + GnRH antagonist (flexible)</td>
<td>0/10 (0)</td>
<td>2/18 (11.1) (1.5 to 28.7)</td>
<td>0.27</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rFSH + GnRH antagonist (start day)</td>
<td>0/10 (0)</td>
<td>1/7 (14.3) (0.4 to 45.9)</td>
<td>0.22</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Rashidi, 2015&lt;sup&gt;143&lt;/sup&gt; RCT (276)</td>
<td>rFSH</td>
<td>hMG</td>
<td>3/25 (12.0) (2.7 to 27.0)</td>
<td>3/18 (16.7) (3.8 to 36.4)</td>
<td>0.63</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Stadtmauer, 2011&lt;sup&gt;159&lt;/sup&gt; RCT (98)</td>
<td>rFSH</td>
<td>rFSH + GnRH antagonist (flexible)</td>
<td>2/12 (16.7) (2.3 to 41.3)</td>
<td>1/19 (5.3) (0.1 to 18.5)</td>
<td>0.30</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>rFSH + GnRH antagonist (start day)</td>
<td>2/12 (16.7) (2.3 to 41.3)</td>
<td>2/9 (22.2) (3.2 to 52.7)</td>
<td>0.75</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS</td>
<td>Rashidi, 2015&lt;sup&gt;143&lt;/sup&gt; RCT (276)</td>
<td>rFSH</td>
<td>hMG</td>
<td>1/132 (0.8) (0.0 to 2.8)</td>
<td>2/144 (1.4) (0.2 to 3.8)</td>
<td>0.61</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; GnRH=gonadotropin-releasing hormone; hMG=human menopausal gonadotropin; N=number of patients; NR=not reported; NS=not statistically significant; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome; RCT=randomized control trial; rFSH=recombinant follicle-stimulating hormone

### 9. Assisted Reproductive Technology (ART) for PCOS

**IVF for PCOS**

Ten RCTs<sup>146,147,150,153,161,162,164,167,171,178</sup> (4 good, 5 fair, 1 poor-quality) compared outcomes for patients with PCOS undergoing IVF. These studies varied in the medication type used for ovulation stimulation, adjunct treatments, and laboratory phase methods. One study explored *in vitro* maturation (IVM) which involves immature oocyte retrieval with subsequent oocyte IVM.<sup>147</sup> These 10 studies included 2,176 patients combined. There was no evidence regarding neonatal outcomes, time to pregnancy, costs, and long-term child or maternal outcomes.

Table 12 summarizes the findings from these studies related to live birth, pregnancy complications, and short-term adverse effects. Live birth was investigated in five studies with one<sup>146</sup> reporting it as any live birth per patient and four<sup>147,150,164,167</sup> reporting any live birth per cycle. One of the studies reporting any live birth/patient compared berberine (an alkaloid extracted from Chinese medicinal herbs), metformin, and placebo as adjunct therapies to IVF. The results were significant for berberine, resulting in more live births than metformin and
placebo; metformin resulted in significantly more live births than placebo. This evidence however was rated as insufficient strength of evidence given findings form one fair quality trial. Of the four studies that reported live birth according to any live birth per cycle, three incorporated gonadotropin use as adjunct therapy. The results of any live birth per cycle were not significant across intervention groups (low SOE).

Three studies examined pregnancy complications, with all three reporting on miscarriage, one reporting on multiple births, and one reporting on ectopic pregnancies. None of the studies reported significant differences between intervention groups for miscarriages (moderate SOE) but given the heterogeneity in intervention protocols we did not perform a meta-analysis of these findings synthesis. Multiple births were reported in one study without significant differences measured but the strength of evidence was rated as insufficient given findings from one fair-quality trial. Ectopic pregnancies were not significantly different in the one fair-quality study that reported this outcome (insufficient SOE).

Five studies examined short-term adverse effects of ART with all 6 reporting on OHSS. One study compared metformin with placebo as oral ovulation induction in addition to IVF. Two studies reported outcomes by severity of OHSS. Heterogeneity in comparisons and findings as well as the studies being underpowered to detect differences in OHSS led to an insufficient strength of evidence rating.

Three observational studies explored the comparative effectiveness of ART with other infertility treatments in women with PCOS. One nationwide birth cohort study identified all pregnancies with a live-single born child over an 8 year period in Denmark and compared the incidence of type I diabetes among those conceived with fertility treatment to those conceived naturally. There was no association between PCOS infertility as an indicator for fertility treatment and the subsequent development of Type I diabetes in offspring (adjusted HR 0.98, 95% CI 0.32 to 3.05) (moderate SOE).

A second Danish national cohort study from registry data examined success rates across complete fertility treatment courses including insemination, ART, and natural conception among couples treated using homologous gametes and no previous live births due to fertility treatment. 13.7% (95% CI 13.1 to 14.4) of women with first treatment by IUI and 4.2% (95%CI 3.7 to 4.6) of women with first treatment with ART were diagnosed with anovulation as the specified cause of female infertility. Across all women, anovulatory infertility was predictor of high live birthrate. The adjusted OR for live birth within 2 years from first IUI treatment was 1.31 (95% CI 1.15 to 1.50; p<0.0001) and 1.57 for ART (95% CI 1.18 to 2.11; p =0.002). Live birthrate differed by maternal age for women with anovulatory infertility less than 35 years 45.3 (95% CI 42.8 to 47.9) with IUI and 14.8 (95% CI 13.1 to 16.7) for ART. This compares to 31.7 (95% CI 25.2 to 38.9) for IUI among women over 35 years and 6.9 (95% CI 3.7 to 11.5) for ART.

Finally, a prospective cohort study of women with PCOS compared maternal and neonatal outcomes with a reference population. Of the 188 included women with PCOS, 14 percent had conceived spontaneously, 68 percent had undergone ovulation induction, and 16 percent underwent IVF/ICSI. A subgroup analysis found no differences in maternal or neonatal (including small for gestational age and neonatal death) complications across presence/absence and type of fertility treatment.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>An, 2014&lt;sup&gt;246&lt;/sup&gt; RCT (109)</td>
<td>Berberine +IVF</td>
<td>Metformin +IVF</td>
<td>18/37 (48.6) (32.9 to 64.5)</td>
<td>14/38 (36.8) (22.5 to 52.5)</td>
<td>0.30</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Placebo +IVF</td>
<td></td>
<td>18/37 (48.6) (32.9 to 64.5)</td>
<td>7/34 (20.6) (9.0 to 35.5)</td>
<td>0.013</td>
<td>Greater live birth with berberine compared to placebo</td>
</tr>
<tr>
<td>Live birth: Any/cycle</td>
<td>Choi, 2012&lt;sup&gt;247&lt;/sup&gt; RCT (61)</td>
<td>IVM + hCG agonist</td>
<td>GnRH antagonist</td>
<td>5/14 (35.7) (13.9 to 61.4)</td>
<td>5/14 (35.7) (13.9 to 61.4)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GnRH antagonist</td>
<td>10/39 (25.6) (13.4 to 40.2)</td>
<td></td>
<td>0.47</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kim, 2012&lt;sup&gt;250&lt;/sup&gt; RCT (208)</td>
<td>GnRH agonist</td>
<td>GnRH antagonist</td>
<td>36/103 (35.0) (26.1 to 44.4)</td>
<td>36/105 (34.3) (25.6 to 43.6)</td>
<td>0.92</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Kurzawa, 2008&lt;sup&gt;252&lt;/sup&gt; RCT (70)</td>
<td>GnRH agonist</td>
<td>GnRH antagonist</td>
<td>18/37 (48.6) (32.9 to 64.5)</td>
<td>14/33 (42.4) (26.4 to 59.4)</td>
<td>0.481</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Ge 2008&lt;sup&gt;257&lt;/sup&gt; RCT (62)</td>
<td>hCG medium</td>
<td>hCG free medium + transfer</td>
<td>9/29 (31.0) (15.9 to 48.7)</td>
<td>10/30 (33.3) (17.9 to 50.8)</td>
<td>0.85</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hCG free medium</td>
<td>10/30 (33.3) (17.9 to 50.8)</td>
<td>2/30 (6.7) (0.8 to 17.8)</td>
<td>0.62</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Kurzawa, 2008&lt;sup&gt;252&lt;/sup&gt; RCT (70)</td>
<td>GnRH agonist +IVF</td>
<td>GnRH antagonist +IVF</td>
<td>5/37 (13.5) (4.7 to 26.1)</td>
<td>3/33 (9.1) (2.0 to 20.8)</td>
<td>0.5</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Kurzawa, 2008&lt;sup&gt;252&lt;/sup&gt; RCT (70)</td>
<td>GnRH agonist +IVF</td>
<td>GnRH antagonist +IVF</td>
<td>2/37 (5.4) (0.7 to 14.5)</td>
<td>6/33 (18.2) (7.2 to 32.8)</td>
<td>0.154</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Ge 2008&lt;sup&gt;257&lt;/sup&gt; RCT (62)</td>
<td>hCG medium</td>
<td>hCG-free medium + transfer</td>
<td>3/29 (10.3) (2.3 to 23.5)</td>
<td>3/30 (10.0) (17.9 to 50.8)</td>
<td>0.97</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>hCG-free medium</td>
<td>3/29 (10.3) (2.3 to 23.5)</td>
<td>2/30 (6.7) (0.8 to 17.8)</td>
<td>0.62</td>
<td>No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------</td>
<td>---------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Wang, 2016</td>
<td>RCT (120)</td>
<td>MPA + hMG+IVF</td>
<td>Short protocol (Decapeptyl, hMG, hCG)</td>
<td>4/75 (5.33)</td>
<td>7/84 (8.33)</td>
<td>0.457</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Wang, 2016</td>
<td>RCT (120)</td>
<td>MPA + hMG+IVF</td>
<td>Short protocol (Decapeptyl, hMG, hCG)</td>
<td>1/49 (2.04)</td>
<td>1/45 (2.22)</td>
<td>0.952</td>
</tr>
<tr>
<td>Palomba, 2011</td>
<td>RCT (120)</td>
<td>Metformin+ IVF</td>
<td>Placebo+IVF</td>
<td>5/60 (8.3) (2.8 to 16.5)</td>
<td>18/60 (30.0) (19.2 to 42.1)</td>
<td>0.003</td>
<td>Reduced OHSS with metformin compared to placebo</td>
</tr>
<tr>
<td>Aboulghar, 2010</td>
<td>RCT (84)</td>
<td>rFSH</td>
<td>uFSH</td>
<td>0/42 (0)</td>
<td>1/42 (2.4) (0.1 to 8.6)</td>
<td>0.31</td>
<td>No difference</td>
</tr>
<tr>
<td>Tehranninejad, 2010</td>
<td>RCT (90)</td>
<td>GnRH agonist</td>
<td>GnRH antagonist</td>
<td>10/45 (22.2) (11.5 to 35.3)</td>
<td>0/45 (0)</td>
<td>0.001</td>
<td>Reduced OHSS with GnRH antagonist compared to GnRH agonist</td>
</tr>
<tr>
<td>Tehranninejad, 2010</td>
<td>RCT (90)</td>
<td>GnRH agonist</td>
<td>GnRH antagonist</td>
<td>5/45 (11.1) (3.8 to 21.7)</td>
<td>0/45 (0)</td>
<td>0.02</td>
<td>Reduced OHSS with GnRH antagonist compared to GnRH agonist</td>
</tr>
<tr>
<td>Aghahosseini, 2017</td>
<td>RCT (100)</td>
<td>Low-dose hCG at time of GnRH agonist</td>
<td>Low-dose hCG 35 hrs after GnRH agonist</td>
<td>13/40 (32.5)</td>
<td>5/40 (12.5)</td>
<td>0.03</td>
<td>Reduced OHSS with hCG given 35 hours after GnRH agonist</td>
</tr>
</tbody>
</table>

*p values for berberine vs. placebo, metformin vs. placebo, and berberine vs. metformin were each <0.05.

Abbreviations: CI=confidence interval; GnRH=gonadotropin-releasing hormone; hCG=human chorionic gonadotropin; hMG=human menopausal gonadotropin; IVF=in vitro fertilization; IVM=in vitro maturation; MPA=medroxyprogesterone acetate; N=number of patients; NR=not reported; NS=not statistically significant; OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome; RCT=randomized control trial; rFSH=recombinant follicle-stimulating hormone; uFSH=urinary follicle-stimulating hormone

### Adjuncts to IVF for PCOS

Three studies looked at adjuncts to IVF in PCOS women. Specifically, two RCTs (one fair quality and one good quality) compared metformin with placebo as pretreatment prior to IVF in PCOS patients. One study presented any live birth per patient for 3 study populations: intention-to-treat (all women randomized to treatment), spontaneous pregnancy (all women with a positive urinary pregnancy test prior to controlled ovarian...
stimulation), and those women randomized to the study who started ovarian stimulation (ART population). Note that we judge only the intention-to-treat results as fair quality, with the others poor (due to increased risk of bias). The other study of pretreatment with metformin reported significantly lower birth rates in the metformin group. The live birth results are presented in Table 13 (insufficient SOE). There was no difference in OHSS in one study177 (Table 13) and no evidence regarding pregnancy complications, neonatal outcomes, time to pregnancy, costs, other short-term adverse effects, and long-term child or maternal outcomes. The third study190 reported on OHSS in women treated with methylprednisolone as an adjunct to IVF. No differences overall or by severity of OHSS were reported. (Table 13). SOE was rated as insufficient for all outcomes given the small studies with varying adjuncts.

Table 13. Outcomes for adjuncts to IVF in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention (Population)</th>
<th>Comparator (Population)</th>
<th>Results Intervention N (%)</th>
<th>Results Comparator N (%)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Kjotrod, 2011155 RCT (149)</td>
<td>Metformin (Intention to treat)</td>
<td>Placebo (Intention to treat)</td>
<td>36/74 (48.6)</td>
<td>24/75 (32.0)</td>
<td>0.038</td>
<td>Greater live births with metformin compared to placebo within the intention to treat analysis</td>
</tr>
<tr>
<td>Live birth: Any/embryo transfer</td>
<td>Jacob, 2016177 RCT (153)</td>
<td>Metformin</td>
<td>Placebo</td>
<td>16/58 (27.6)</td>
<td>33/64 (51.6)</td>
<td>0.01</td>
<td>Lower birth rates in metformin group.</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS, any</td>
<td>Jacob, 2016177 RCT (153)</td>
<td>Metformin</td>
<td>Placebo</td>
<td>21/75 (28.0)</td>
<td>16/74 (21.6)</td>
<td>0.726</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Yeganeh, 2018190 RCT (219)</td>
<td>Methylprednisolone</td>
<td>Placebo</td>
<td>18/93 (19.4)</td>
<td>15/91 (16.5)</td>
<td>0.61</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS, mild</td>
<td>Yeganeh, 2018190 RCT (219)</td>
<td>Methylprednisolone</td>
<td>Placebo</td>
<td>10/93 (10.8)</td>
<td>5/91 (5.5)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS, moderate/severe</td>
<td>Jacob, 2016177 RCT (153)</td>
<td>Metformin</td>
<td>Placebo</td>
<td>12/75 (16)</td>
<td>9/74 (12.2)</td>
<td>0.66</td>
<td>No difference</td>
</tr>
</tbody>
</table>
### Fresh Versus Frozen Embryos in IVF for PCOS

One good-quality RCT\(^{174}\) compared outcomes in women with PCOS undergoing IVF with transfer of frozen versus fresh embryos. Outcomes presented in Table 14 include live births, multiple births, and parental and neonatal outcomes. Live birth rates were significantly higher in the frozen embryo transfer group, with no significant difference in multiple births. Miscarriages and ectopic pregnancies were higher in the fresh embryo transfer group. No significant differences were observed for neonatal deaths or congenital abnormalities. SOE was low for all outcomes given findings from one study.

#### Table 14. Outcomes for fresh versus frozen embryo transfer in IVF in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Chen, 2016(^{174}) RCT (1508)</td>
<td>Frozen Embryo Transfer in IVF</td>
<td>Fresh Embryo Transfer in IVF</td>
<td>368/746 (49.3)</td>
<td>320/762 (42.0)</td>
<td>0.004</td>
<td>Greater live births with frozen embryo transfer</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Chen, 2016(^{174}) RCT (1508)</td>
<td>Frozen Embryo Transfer in IVF</td>
<td>Fresh Embryo Transfer in IVF</td>
<td>118/746 (15.8)</td>
<td>108/762 (14.2)</td>
<td>0.41</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Chen, 2016(^{174}) RCT (1508)</td>
<td>Frozen Embryo Transfer in IVF</td>
<td>Fresh Embryo Transfer in IVF</td>
<td>108/492 (22.0)</td>
<td>161/492 (32.7)</td>
<td>&lt;0.001</td>
<td>Lower miscarriages with frozen embryo transfer</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Chen, 2016(^{174}) RCT (1508)</td>
<td>Frozen Embryo Transfer in IVF</td>
<td>Fresh Embryo Transfer in IVF</td>
<td>10/492 (2.0)</td>
<td>54/762 (7.1)</td>
<td>&lt;0.001</td>
<td>Lower ectopic pregnancies with frozen embryo transfer</td>
</tr>
</tbody>
</table>

Abbreviations: ART=assisted reproductive technology; CI=confidence interval, N=number of patients; NS=not statistically significant; OHSS=ovarian hyperstimulation Syndrome; PCOS=polycystic ovary syndrome; RCT=randomized control trial
### Table 15. Outcomes for comparisons of ICSI treatment in women with PCOS

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Live birth:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any/patient</td>
<td>Zheng, 2012(^{149})</td>
<td>hCG Priming</td>
<td>No hCG Priming</td>
<td>9/40 (22.5) (11.1 to 36.5)</td>
<td>13/42 (31.0) (18.1 to 45.5)</td>
<td>0.39</td>
<td>No difference</td>
</tr>
<tr>
<td>RCT (82)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pregnancy complications:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Miscarriage</td>
<td>Hosseini, 2010(^{158})</td>
<td>GnRH agonist</td>
<td>GnRH antagonist</td>
<td>4/55 (7.3) (2.1 to 15.4)</td>
<td>8/57 (17.6) (6.9 to 25.8)</td>
<td>0.219</td>
<td>No difference</td>
</tr>
<tr>
<td>RCT (112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval, IVF=*in vitro* fertilization; N=number of patients; PCOS=polycystic ovary syndrome; RCT=randomized control trial

**ICSI for PCOS**

Two fair-quality RCTs\(^{149,158}\) compared different methods for ICSI. These studies varied in the medication type used for ovulation stimulation and adjunct treatments. These 2 studies included 186 patients combined. Outcomes are summarized in Table 15. Live birth was investigated in one study,\(^{149}\) reporting it as any live birth per patient. The pregnancy complication of miscarriage was investigated in the other study.\(^{158}\) Neither outcome was significantly different between treatments but the strength of evidence was rated as insufficient given findings from one small fair-quality study. Only the second study evaluated the short-term adverse effect of OHSS. The use of GnRH antagonist reduced the incidence of OHSS significantly, primarily through its effect on moderate/severe OHSS although again the evidence from one fair-quality study results in an insufficient strength of evidence rating.

There was no evidence regarding neonatal outcomes, time to pregnancy, costs, and long-term child or maternal outcomes.

We also identified one good-quality meta-analysis that examined HCG priming for fertility treatments with *in vitro* maturation.\(^{196}\) For women with PCOS, outcomes of interest were reported in only one study\(^{149}\) which we have described in our results of individual studies.
### Subgroups of Interest for PCOS

Three good-quality RCTs\(^{128,151,160}\) reported live birth with BMI analyzed as a subgroup in women with PCOS. Interventions and BMI parameters varied between studies. One study\(^{151}\) reported BMI as less than 27 or \(\geq 27\). This study found no significant differences in live births between metformin and placebo as adjunct therapies to clomiphene by BMI subgroups (BMI <27 45.5% vs. 34.2%; \(p=0.16\) | BMI \(\geq 27\) 35.9% vs. 22.6%; \(p=0.1\)).

Another study\(^{160}\) reported findings for women with BMI \(\leq 32\) or >32. This study compared clomiphene, metformin, and clomiphene plus metformin. Pregnancy and live birth rates were low in women with BMI > 32 whatever treatment was used, with no evidence of benefit of metformin over placebo. For women with BMI \(\leq 32\) there was no evidence of significant differences in outcomes whether treated with metformin, clomiphene or both.

The third study\(^{128,129}\) compared clomiphene and clomiphene plus metformin (combination therapy) in 626 women. Metformin was used as a reference in the analysis. Patients were divided into BMI subgroups (<30, 30-34, \(\geq 35\)). When the treatment arms were stratified by BMI, both clomiphene and combination therapy were more successful regarding the outcome of live birth compared to metformin with greatest benefits seen in the low and high BMI categories.

These findings suggest that the addition of metformin to clomiphene does not significantly improve live birth weight in obese women with PCOS. Studies however were not powered to detect differences within subgroups, and different categorization definitions for BMI limit the ability to pool study results.

### Strength of Evidence for PCOS

Table 16 summarizes the SOE for the findings described above. In general, SOE was judged insufficient or low for most outcomes with a few exceptions: There was moderate SOE live births with the use of letrozole versus clomiphene, and for oral agents versus surgical management. There was moderate SOE for miscarriage rates between clomiphene versus...
metformin and for oral agents versus surgical management. A common limitation across all comparisons was lack of precision for estimates of rare but important harms such as OHSS or surgical complications

Table 16. Strength of evidence for major outcomes—KQ 1 (PCOS)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>SOE (Rationale)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral agents alone: Letrozole vs. Berberine vs. Berberine + Letrozole</td>
<td>Live birth (any/patient)</td>
<td>1 RCT&lt;sup&gt;175&lt;/sup&gt; (644)</td>
<td>Improvement: Letrozole or letrozole and berberine increase live birth rates compared to berberine alone.</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;175&lt;/sup&gt; (644)</td>
<td>No difference: No significant difference between letrozole, berberine, or combination therapy</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;175&lt;/sup&gt; (644)</td>
<td>No difference: No significant difference between letrozole, berberine, or combination therapy</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT&lt;sup&gt;175&lt;/sup&gt; (644)</td>
<td>No difference: No significant difference between letrozole, berberine, or combination therapy</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td>Oral agents alone: Letrozole vs. Clomiphene</td>
<td>Live birth (any/patient)</td>
<td>2 RCTs&lt;sup&gt;131,179&lt;/sup&gt; (909)</td>
<td>Improvement: Letrozole has higher live birth rates than clomiphene (pooled OR 1.64; 95% CI, 1.32 to 2.04).</td>
<td>Moderate (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>3 RCTs&lt;sup&gt;131,179&lt;/sup&gt; (886)</td>
<td>Improvement: Letrozole has lower rates of multiple birth compared to clomiphene (OR 0.38; 95% CI, 0.17 to 0.84)</td>
<td>Moderate (Inconsistent)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>3 RCTs&lt;sup&gt;131,179&lt;/sup&gt; (886)</td>
<td>No difference: No difference between letrozole and clomiphene (OR 0.72, 95% CI 0.24 to 2.17).</td>
<td>Moderate (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs&lt;sup&gt;131,179&lt;/sup&gt; (886)</td>
<td>No difference: No statistical difference between letrozole and clomiphene (pooled OR 1.32, 95% CI 0.92 to 1.88)</td>
<td>Moderate (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Death</td>
<td>1 RCT&lt;sup&gt;131&lt;/sup&gt; (750)</td>
<td>Inconclusive: Given the rarity of the outcome, much larger data sets are needed.</td>
<td>Insufficient (Imprecise, findings from only 1 study, small number of events)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT&lt;sup&gt;131&lt;/sup&gt; (750)</td>
<td>No difference: No significant difference in birthweight between letrozole and clomiphene</td>
<td>Low (Findings from only 1 study)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>2 RCTs&lt;sup&gt;131,179&lt;/sup&gt; (909)</td>
<td>Inconclusive: No significant difference in congenital anomaly rates between letrozole and clomiphene However, given the rarity of the outcomes, much larger data sets are needed.</td>
<td>Insufficient (Imprecise, small number of events)</td>
</tr>
<tr>
<td></td>
<td>Time to pregnancy</td>
<td>1 RCT&lt;sup&gt;131&lt;/sup&gt; (750)</td>
<td>No difference: No significant difference in time to pregnancy between clomiphene vs. letrozole</td>
<td>Low (Findings from only 1 study)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>SOE (Rationale)*</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Oral agents alone: Clomiphene vs. Metformin vs. Metformin + Clomiphene</td>
<td>Live birth (any/patient)</td>
<td>3 RCTs [128,151,173] (824)</td>
<td>No difference: No statistical difference between clomiphene and metformin or between clomiphene and combination therapy of metformin and clomiphene (OR 1.21, 95% CI 0.92 to 1.59)</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 SRs (3 studies, 912 patients(^3)); (9 studies, 1079 patients(^3))</td>
<td>No difference: No differences in multiple birth rates between clomiphene alone, metformin alone, and clomiphene plus metformin</td>
<td>Low (Imprecise, suspected reporting bias)</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td></td>
<td>3 RCTs [128,160,166] (921)</td>
<td>No difference: No difference between studied oral agents. Very few ectopic pregnancies overall.</td>
<td>Low (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 SR(^{199}) (9 studies, 1079 patients)</td>
<td>Increase: Higher rate of miscarriage in the combined therapy group (clomiphene and metformin) compared to clomiphene alone (OR 1.59, 95% CI, 1.03 to 2.46)</td>
<td>Low (Suspected reporting bias, imprecise)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td></td>
<td>3 RCTs [160,166,173] (1,005)</td>
<td>Inconclusive: SOE was insufficient given the imprecise evidence from identified studies. Given the rarity of the outcomes, much larger data sets are needed.</td>
<td>Insufficient (Imprecise, suspected reporting bias)</td>
</tr>
<tr>
<td>Neocnata outcomes: Congenital anomalies</td>
<td></td>
<td>3 RCTs [160,166,173] (1,005)</td>
<td>No difference: No significant difference in time to pregnancy between clomiphene vs. metformin</td>
<td>Low (Findings from 1 study)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 RCT(^{151}) (343)</td>
<td>Inconclusive: SOE was insufficient given the findings from 1 small study with potential limitations</td>
<td>Insufficient (1 study, moderate limitations)</td>
</tr>
<tr>
<td>Oral agents alone: Clomiphene vs. Tamoxifen</td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT(^{192}) (130)</td>
<td>Inconclusive: SOE was insufficient given the findings from 1 small study with potential limitations</td>
<td>Insufficient (1 study, moderate limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT(^{192}) (130)</td>
<td>Inconclusive: SOE was insufficient given the findings from 1 small study with potential limitations</td>
<td>Insufficient (1 study, moderate limitations)</td>
</tr>
<tr>
<td>Oral agents alone: Clomiphene vs. Tamoxifen</td>
<td>Live birth (any/patient)</td>
<td>1 RCT(^{191}) (88)</td>
<td>No difference: No significant difference in live birth rates between tamoxifen and clomiphene (OR 1.24, 95% CI 0.59-2.62)</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 SR(^{198}) (2 studies, 195 women)</td>
<td>No difference: No significant difference in miscarriage rates between tamoxifen and clomiphene (OR 1.81, 95% CI 0.80-4.12)</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Oral agents alone vs. acupuncture: Active</td>
<td>Live birth</td>
<td>1 RCT(^{188}) (1000)</td>
<td>Improvement: Live birth rates significantly higher for clomiphene vs. placebo; not significantly different for active vs. control acupuncture</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
</tbody>
</table>
| Comparison | Outcome | Study Design (Sample Size) | Conclusion | SOE (Rationale)  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Acupuncture + Clomiphene vs. Control</td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT&lt;sup&gt;188&lt;/sup&gt; (1000)</td>
<td><strong>No difference</strong>: no significant difference in ectopic pregnancy rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td>Acupuncture + Clomiphene vs. Active Acupuncture + Placebo vs. Control Acupuncture + Placebo</td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;188&lt;/sup&gt; (1000)</td>
<td><strong>No difference</strong>: no significant difference in miscarriage rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Congenital Abnormalities</td>
<td>1 RCT&lt;sup&gt;188&lt;/sup&gt; (1000)</td>
<td><strong>No difference</strong>: no significant difference in congenital abnormality rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td></td>
<td>Neonatal Death</td>
<td>1 RCT&lt;sup&gt;188&lt;/sup&gt; (1000)</td>
<td><strong>No difference</strong>: no significant difference in neonatal death rates between oral agents and acupuncture strategies.</td>
<td>Low (1 study with potential risk of bias)</td>
</tr>
<tr>
<td>Oral agents alone vs. oral agents with IUI: Clomiphene without IUI vs. Clomiphene with IUI</td>
<td>Live birth</td>
<td>1 RCT&lt;sup&gt;157&lt;/sup&gt; (188)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;157&lt;/sup&gt; (188)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT&lt;sup&gt;157&lt;/sup&gt; (188)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;157&lt;/sup&gt; (188)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td>Oral agents alone vs. oral agents with IUI: Myo-inositol + Ovulation induction + IUI vs. Ovulation induction and IUI alone</td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;182&lt;/sup&gt; (196)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td></td>
<td>Adverse events: OHSS</td>
<td>1 RCT&lt;sup&gt;182&lt;/sup&gt; (196)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given evidence from 1 small trial.</td>
<td>Insufficient (Imprecise, 1 small trial)</td>
</tr>
<tr>
<td>Oral agents vs. LOD</td>
<td>Live birth (any/patient)</td>
<td>1 SR&lt;sup&gt;85&lt;/sup&gt; (8 studies, 1,034 women)</td>
<td><strong>No difference</strong>: No statistically significant differences between LOD and oral agents (OR = 0.77, 95% CI 0.59 to 1.01)</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 SR&lt;sup&gt;85&lt;/sup&gt; (15 studies, 1,129 women)</td>
<td><strong>Reduction</strong>: There was a reduction in multiple births given LOD as compared to oral agents (OR 0.21, 95% CI 0.08 to 0.58)</td>
<td>Moderate (Suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;189&lt;/sup&gt; (80)</td>
<td><strong>No difference</strong>: No significant differences in miscarriage between LOD and oral agents (OR 1.1, 95% CI 0.74 to 1.61)</td>
<td>Low (Imprecise, suspected reporting bias)</td>
</tr>
<tr>
<td></td>
<td>1 SR&lt;sup&gt;85&lt;/sup&gt; (15 studies, 1,592 women)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 RCTs&lt;sup&gt;165,168&lt;/sup&gt; (328)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given the imprecision and small numbers of events in the identified studies.</td>
<td>Insufficient (Imprecise, suspected reporting bias, 1 small study)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>SOE (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------------------------</td>
<td>------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Short term adverse effects of treatment: Surgical complications</td>
<td><strong>Conclusion</strong>: SOE was insufficient given the imprecision and small numbers of events in the identified studies.</td>
<td><strong>SOE (Rationale)</strong>: Insufficient (Imprecise, suspected reporting bias, small sample size)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Oral agents alone vs. gonadotropins: Clomiphene citrate vs. low-dose FSH</strong></td>
<td><strong>Live birth</strong>: Inconclusive: SOE was insufficient given the inconsistent evidence from included studies</td>
<td><strong>Insufficient (Imprecise, inconsistent)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td><strong>1 RCT</strong>&lt;sup&gt;152&lt;/sup&gt; (302)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 identified trial.</td>
<td><strong>Insufficient (Imprecise, 1 study)</strong></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td><strong>3 RCTs</strong>&lt;sup&gt;152,176,187&lt;/sup&gt; (1072)</td>
<td><strong>No difference</strong>: Ectopic pregnancy rate did not differ between FSH and clomiphene strategies.</td>
<td><strong>Low (Imprecise)</strong></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td><strong>3 RCTs</strong>&lt;sup&gt;152,176,187&lt;/sup&gt; (1072)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given the inconsistent evidence from included studies</td>
<td><strong>Insufficient (Imprecise, inconsistent)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Oral agents alone vs. gonadotropins Clomiphene plus urinary FSH vs. FSH alone</strong></td>
<td><strong>Live birth</strong>: Inconclusive SOE was insufficient given only 1 identified trial</td>
<td><strong>Insufficient (Imprecise, 1 small study)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifestyle modifications + IVF vs. IVF alone</td>
<td><strong>Live birth</strong>: No difference: No difference in live birth rates for women who underwent lifestyle modification in combination with IVF compared with IVF alone</td>
<td><strong>Moderate (Heterogeneity in interventions)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>LEC vs. rFSH LEC vs. hMG</strong></td>
<td><strong>Live birth</strong>: Inconclusive: SOE was insufficient given only 1 study with moderate risk of bias.</td>
<td><strong>Insufficient (Imprecise, 1 study with limitations)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td><strong>1 RCT</strong>&lt;sup&gt;141&lt;/sup&gt; (168)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 study with moderate risk of bias.</td>
<td><strong>Insufficient (Imprecise, 1 study with limitations)</strong></td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td><strong>1 RCT</strong>&lt;sup&gt;141&lt;/sup&gt; (168)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 study with moderate risk of bias.</td>
<td><strong>Insufficient (Imprecise, 1 study with limitations)</strong></td>
<td></td>
</tr>
<tr>
<td>Short term adverse effects of treatment: OHSS</td>
<td><strong>1 RCT</strong>&lt;sup&gt;148&lt;/sup&gt; (104)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 study with moderate risk of bias.</td>
<td><strong>Insufficient (Imprecise, 1 study with limitations)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>rFSH vs. rFSH + GnRH antagonist or hMG</strong></td>
<td><strong>Live birth (any/cycle)</strong>: Inconclusive: SOE was insufficient given studies with moderate risk of bias and varying definitions of live birth</td>
<td><strong>Insufficient (Moderate study limitations)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td><strong>2 RCTs</strong>&lt;sup&gt;143,159&lt;/sup&gt; (374)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given imprecise findings from studies of moderate risk of bias.</td>
<td><strong>Insufficient (Imprecise findings with moderate study limitations)</strong></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>SOE (Rationale)</td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------------------------------</td>
<td>----------------------------</td>
<td>------------------------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>2 RCTs&lt;sup&gt;144,159&lt;/sup&gt; (374)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given imprecise findings from studies of moderate risk of bias.</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Short term adverse effects of treatment: OHSS</td>
<td>1 RCT&lt;sup&gt;143&lt;/sup&gt; (276)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only 1 study with moderate risk of bias.</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>ART IVF: GnRH agonist +/- IVF vs. GnRH antagonist +/- IVF</td>
<td>Live birth (patient)</td>
<td>1 RCT&lt;sup&gt;146&lt;/sup&gt; (109)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only 1 study with moderate risk of bias and with imprecise findings</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Live birth (cycle)</td>
<td>4 RCTs&lt;sup&gt;145,150,164,167&lt;/sup&gt; (408)</td>
<td><strong>No difference:</strong> No significant difference in included studies but varying interventions and comparators with low numbers of live birth</td>
<td>Low (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;164&lt;/sup&gt; (70)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT&lt;sup&gt;171&lt;/sup&gt; (120)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs&lt;sup&gt;164,167,171&lt;/sup&gt; (279)</td>
<td><strong>No difference:</strong> No differences in miscarriage rates for GnRH agonist vs. antagonist, or hCG medium, hCG-free medium with transfer, and hCG-free medium without transfer.</td>
<td>Moderate (Imprecise findings with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Short term adverse effects of treatment: OHSS</td>
<td>5 RCTs&lt;sup&gt;153,161,162,164,178&lt;/sup&gt; (468)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given that all identified studies were underpowered to detect differences in OHSS.</td>
<td>Insufficient (Imprecise, underpowered studies)</td>
<td></td>
</tr>
<tr>
<td>Adjunct to IVF</td>
<td>Live birth (patient)</td>
<td>1 RCT&lt;sup&gt;155&lt;/sup&gt; (149)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 RCTs&lt;sup&gt;177,190&lt;/sup&gt; (372)</td>
<td><strong>Inconclusive:</strong> SOE was insufficient given only small studies with moderate risk of bias and varying adjuncts.</td>
<td>Insufficient (Moderate study limitations)</td>
</tr>
<tr>
<td>Fresh vs. Frozen Embryos in IVF for PCOS</td>
<td>Live birth (any/cycle)</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>Improvement:</strong> Live birth rates were significantly higher with frozen embryo transfer compared to fresh embryos</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>No difference:</strong> No difference in multiple births with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>Reduction:</strong> Ectopic pregnancies were reduced with frozen embryo transfer</td>
<td>Low (1 study)</td>
</tr>
</tbody>
</table>
### Comparison Table

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>SOE (Rationale)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>Reduction</strong>: Miscarriages were reduced with frozen embryo transfer</td>
<td>Low (1 study)</td>
<td></td>
</tr>
<tr>
<td>Neonatal Outcomes: Congenital abnormalities</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>No difference</strong>: No difference congenital abnormalities with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
<td></td>
</tr>
<tr>
<td>Neonatal Death</td>
<td>1 RCT&lt;sup&gt;174&lt;/sup&gt; (1508)</td>
<td><strong>No difference</strong>: No difference neonatal deaths with fresh versus frozen embryo transfer</td>
<td>Low (1 study)</td>
<td></td>
</tr>
<tr>
<td>ART vs. no infertility treatment</td>
<td>Long-term outcomes: Child (type 1 diabetes mellitus)</td>
<td>1 Obs&lt;sup&gt;164&lt;/sup&gt; (565,116 pregnancies)</td>
<td><strong>No difference</strong>: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with PCOS infertility conceived with ART compared to children conceived with no fertility treatment</td>
<td>Moderate (Imprecise)</td>
</tr>
<tr>
<td>ART ICSI: GnRH agonist vs. GnRH antagonist</td>
<td>Live birth</td>
<td>1 RCT&lt;sup&gt;149&lt;/sup&gt; (82)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Imprecise findings with moderate study limitations, 1 small study)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;158&lt;/sup&gt; (112)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Imprecise findings with moderate study limitations, 1 small study)</td>
<td></td>
</tr>
<tr>
<td>Short term adverse effects of treatment: OHSS</td>
<td>1 RCT&lt;sup&gt;158&lt;/sup&gt; (112)</td>
<td><strong>Inconclusive</strong>: SOE was insufficient given only 1 small study with moderate risk of bias.</td>
<td>Insufficient (Imprecise findings with moderate study limitations, 1 small study)</td>
<td></td>
</tr>
</tbody>
</table>

*Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

**Abbreviations:** ART=assisted reproductive technology; BMI=body mass index; CI=confidence interval; FSH=follicle-stimulating hormone; GnRH=gonadotropin-releasing hormone; hCG=human chorionic gonadotropin; hMG=human menopausal gonadotropin; IUI=intrauterine insemination; IVF=intracytoplasmic sperm injection; IVM=in vitro maturation; KQ=Key Question; LEC=laparoscopic electrocauterization; LOD=laparoscopic ovarian drilling/diathermy; N=number of patients/participants; NA=not applicable; NR=not reported; Obs=observational study; OHSS=ovarian hyperstimulation syndrome; OR=odds ratio; PCOS=polycystic ovary syndrome; RCT=randomized controlled trial; rFSH=recombinant follicle-stimulating hormone; SR=systematic review; uFSH=urinary follicle-stimulating hormone.

**Key Question 2: Endometriosis**

**KQ 2.** What are the comparative safety and effectiveness of available treatment strategies for women with endometriosis who are infertile and who wish to become pregnant?

**KQ 2a.** Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, stage of endometriosis, presence of other potential causes of female infertility, or presence of male factor infertility?
Endometriosis is a condition defined by the presence of endometrial tissue outside the uterus, most commonly on the ovary and peritoneum. It is observed in approximately 6–10 percent of women of fertile age, and in up to 35–50 percent of women with infertility, pelvic pain, or both. Endometriosis severity is defined in a system of stages that are based on a weighted point system. The severity of the disease is classified as minimal (Stage I), mild (Stage II), moderate (Stage III), or severe (Stage IV) based on surgical findings.

**Description of Included Studies for KQ 2 (Endometriosis)**

We identified 7 individual studies that addressed infertility treatment for women with endometriosis. Three studies were RCTs; one was good quality, and two were fair quality. The remaining four were observational cohorts; two were good quality, and two were fair quality. Geographical locations varied; one study was located in Africa, one was located in Asia, one in Australia/NZ, two in the United States, and the remaining two were located in the United Kingdom or continental Europe. Five studies were conducted in subspecialty practices. One study reported funding from government sources, one study reported funding from industry, one study reported non-government, non-industry funding, and the remaining four studies did not report a funding source or the source was unclear.

The main classes of treatment examined were: oral ovulation/superovulation with or without IUI, gonadotropins with or without IUI, ART (IVF or ICSI), surgical treatment, and surgery with hormonal adjunctive therapy. We did not perform meta-analysis because of the lack of studies reporting results for similar outcomes and treatment comparisons.

In addition to the above studies, we discuss two good-quality systematic reviews that addressed the comparative effectiveness of various treatments for infertility in women with endometriosis, and incorporate the consistency of their findings with included studies into our strength of evidence (SOE) ratings.

**Key Points for Endometriosis**

Key findings for couples where the primary cause of infertility is endometriosis in the female partner included:

- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)
- The live birth rate per cycle was higher in couples who underwent ART than those who used IUI (low SOE)
- SOE was rated insufficient for all other comparisons/outcomes.

**Detailed Synthesis by Treatment for Endometriosis**

Included studies and their findings for the following treatments for endometriosis are detailed in this section:

1. Oral Agents With IUI
2. Adjunctive Hormonal Therapy After Surgical Treatment
3. ART
1. Oral Agents With IUI After Laparoscopic Treatment for Endometriosis

One good-quality RCT examined pregnancy outcomes after ovulation induction with IUI in 136 infertile women with minimal/mild endometriosis who sought infertility treatment 6 to 12 months after laparoscopic treatment. Women were randomized to treatment with clomiphene or letrozole for 5 days starting on day 3 of the menstrual cycle. Human chorionic gonadotropin (hCG; 10,000 IU) was given intramuscularly when one follicle measuring at least 18 mm was identified. IUI was performed 32–36 hours after HCG injection. There was no significant difference in live birth rate among women treated with letrozole and IUI (44.9%) as compared with clomiphene citrate and IUI (40.3%; RR 0.89; 95% CI, 0.43 to 1.58) (low SOE). There was also no significant difference in miscarriage rate among women treated with letrozole and IUI (12.4%) as compared with clomiphene citrate and IUI (12.9%; RR 1.13; 95% CI, 0.76 to 1.68). Both outcomes were rated as insufficient SOE given findings from one small study. Other pregnancy complication outcomes, neonatal outcomes, time to pregnancy, costs, short term adverse effects of treatments, or long term effects were not evaluated.

2. Adjunctive Hormonal Therapy After Surgical Treatment for Endometriosis

Two fair-quality RCTs evaluated the use of hormonal therapy after surgery for IVF. A fair-quality RCT examined the efficacy of hormonal therapy as an adjunctive treatment with surgical correction of endometriosis. This trial studied short-term oral contraceptives, with and without a “Dan’e” Chinese herbal mixture, compared with no treatment, immediately after surgery in 156 women with endometriosis. After the surgical treatment, participants were randomly allocated to three groups: (a) a combined oral contraceptive for 63 days, (b) combined oral contraceptive for 63 days + Dan’e Chinese herbal mixture for the last 30 days, or (c) no medical treatment. The mean duration of follow-up was 22.17 ± 3.39 months.

A total of 73 pregnancies occurred. The total pregnancy rate was 46.80%. There was no difference in live birth rate (70% vs. 81.2% vs. 79.2%), ectopic pregnancy rate (5% vs. 0 vs. 8.3%), or miscarriage rate (20% vs. 18.7% vs. 12.5%) in any of the three arms of the study (p >0.05). One stillbirth occurred in this study, in the group that was treated with oral contraceptive alone. Other outcomes of interest were not evaluated.

A second fair-quality RCT, performed in Russia, examined 144 women who had laparoscopic surgery for removal of ovarian endometriomas and tested the effects of 3 agents for 6 months prior to IVF. The 3 treatments tested were 1) Progestin, dienogest 2mg orally daily; 2) GnRH agonist, triptorelin 3.75 mg intramuscularly monthly; 3) no hormonal therapy prior to IVF. The live birth rates for participants who received dienogest were significantly higher compared to those on no therapy with live birth rates of 36.8% versus 11.1% respectively, P=0.013. The live birth rates for those receiving the GnRH agonist compared to no therapy were not statistically significantly different and were 28.6% versus 11.1% respectively, P=0.234.

The imprecise evidence from fair-quality studies performed in Russia and China resulted in insufficient strength of evidence across the outcomes assessed.

3. ART for Endometriosis

Two observational studies (1 fair, 1 good quality) looked at the use of ART interventions in women with endometriosis-related infertility.
A good-quality cohort study of the national Danish ART registry and Medical Birth registry compared live birth rates after treatment with ART versus IUI for women with endometriosis, stratified by age.\(^{186}\) Compared to IUI and natural conception, live birth rates were higher among women who received ART. For women ≥ 35 years of age live birthrate (95% CI) for women undergoing ART was 39.0 (24.2, 55.5); undergoing IUI 0.0 (0.0 to 8.6); and natural conception 4.9 (0.6 to 16.5).\(^{186}\) For women <35 years of age, live birthrate (95% CI) for women undergoing ART was 51.3 (44.1 to 58.4); undergoing IUI 0.5 (0.0 to 2.8); and natural conception 9.6 (5.9 to 14.7) (low SOE).

One fair-quality observational study\(^{123}\) examined the effectiveness of ART using data from the Society for Assisted Reproductive Technology Clinic Outcome Reporting System (SART CORS) database, a collective cohort of 305,774 pregnancies in the United States. The good-quality retrospective analysis examined live birth rates, maternal complications, and neonatal complications based on the number of embryos transferred in 69,028 ART cycles.\(^{123}\) In this study, 7,104 cycles of IVF or ICSI were performed in couples with the diagnosis of endometriosis. In this endometriosis subgroup, the live birth rate per cycle after treatment with either IVF or ICSI was 48.7%. For women with endometriosis, the live birth rate per cycle was higher in couples who underwent two embryo transfer (n=3808 cycles, live birth rate=51.5%) as compared with single-embryo transfer (n=319 cycles, live birth rate=46.6%) (p<0.0001) (low SOE).

**Subpopulations of Interest for Endometriosis**

A 2014 good-quality systematic review compared IVF outcomes in 20,167 women (78 publications) with endometriosis (2227 women with stage I-II endometriosis, 1703 women with stage III-IV endometriosis) to 121,931 women without endometriosis.\(^{64}\) This publication included randomized and nonrandomized controlled trials. They concluded that live birth rates were similar regardless of whether surgical correction of endometrioma was performed. Women with more severe endometriosis had similar outcomes as compared with women with less severe endometriosis. Specifically, for the comparison of women with Stage-III/IV versus Stage-I/II endometriosis, results were as follows: live birth, RR 0.94 (95% CI, 0.80 to 1.11); clinical pregnancy, RR 0.90 (95% CI, 0.82 to 1.00); miscarriage, RR 0.99 (95% CI, 0.73 to 1.36).

In 2015, Hamdan and colleagues published a good-quality systematic review\(^{60}\) investigating the association between endometriosis and reproductive outcomes in women undergoing ART. This meta-analysis included 33 studies (30 retrospective observational, 3 RCTs). Fourteen of these studies overlapped with systematic review by Barbosa and colleagues discussed above.\(^{64}\) In women with endometrioma, those who had surgical treatment before IVF/ICSI had a similar live birth rate (OR 0.90; 95% CI, 0.63 to 1.28), a similar clinical pregnancy rate (OR 0.97; 95% CI, 0.78 to 1.20), and a similar miscarriage rate (OR 1.32; 95% CI, 0.66 to 2.65) compared to those with untreated endometrioma.

In summary, the severity of endometriosis, surgical treatment of endometriosis prior to ART, and adjunctive hormonal therapy do not appear to influence fertility outcomes.\(^{60,64}\)

**Strength of Evidence for Endometriosis**

Table 17 summarizes the SOE for the findings described above. In general, the SOE across all outcomes was judged to be insufficient or low, primarily due to imprecision and small numbers of studies, especially for both short-term harms (such as OHSS) and long-term harms.
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Letrozole vs. Clomiphene with IUI</td>
<td>Live birth (any/patient)</td>
<td>1 RCT(^{204}) (136)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings from one small study with limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td></td>
<td>1 RCT(^{204}) (136)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings from one small study with limitations)</td>
</tr>
<tr>
<td>Hormonal Therapy After Surgical Treatment vs. no treatment</td>
<td>Live birth</td>
<td>2 RCTs(^{202,203}) (300)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 2 studies from non-US settings with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td></td>
<td>1 RCT(^{202}) (156)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td></td>
<td>1 RCT(^{202}) (156)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Neonatal outcomes: Death</td>
<td></td>
<td>1 RCT(^{202}) (156)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>ART: IVF/ICSI vs. no treatment</td>
<td>Live birth</td>
<td>1 Obs(^{23}) (69,028 cycles)</td>
<td>Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent 2 embryo transfer (51.5%) as compared with single-embryo transfer (46.6%) (p&lt;0.0001).</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>IUI vs. ART</td>
<td>Live birth</td>
<td>1 Obs(^{186}) (19,884)</td>
<td>Improvement: For women with endometriosis, the live birth rate per cycle was higher in couples who underwent ART than those who used IUI</td>
<td>Low (one study)</td>
</tr>
</tbody>
</table>

\(^a\)Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: ART=assisted reproductive technology; CI=confidence interval; HR=hazard ratio; ICSI=intracytoplasmic sperm injection; IUI=intruterine insemination; IVF= *in vitro* fertilization; KQ=Key Question; N=number of patients/participants; NA=not applicable; Obs=observational study; RCT=randomized controlled trial; RR=relative risk
Key Question 3. Unexplained Infertility

KQ 3. What are the comparative safety and effectiveness of available treatment strategies for women who are infertile for unknown reasons and who wish to become pregnant?

KQ 3a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, presence of other potential causes of female infertility, or presence of male factor infertility?

Unexplained infertility is defined as infertility with no other documented female or male diagnosis.

Description of Included Studies for KQ 3 (Unexplained Infertility)

We identified 50 individual studies that met inclusion criteria for KQ 3.52,119,122,135,139,185,186,206-247

Twenty-one studies were conducted in the United Kingdom or continental Europe,119,135,139,186,206,207,209,211-213,217,218,221,224-226,230,237,242,244 10 in the United States,52,122,123,185,210,213,223,233,236,239 11 in the Middle East,208,216,220,222,229,231,232,241,243,245,246 three in Asia,214,219,247 four in Africa,227,228,234,235 and one in Australia/NZ.238 Forty-two studies were RCTs; 18 were good quality,52,119,135,139,208-210,226-230,233-235,237,238,246 22 were fair quality,206,207,212,214,216-225,231,232,240,242-245,247 and 2 were poor quality.211,241 Eight studies were observational studies; four were good quality,186,215,236,239 and four were fair quality.122,123,185,213 Eleven studies were solely government-funded,52,119,122,135,139,208,210,226-230,233-235,237,238,246 seven studies received non-government, non-industry funding,123,211,215,235,241,244,247 four were solely funded by industry,186,212,220,223 and one study reported a combination of funding sources.236 The remaining 27 studies did not report a funding source or the funding source was unclear. The type of practice was not specified in six studies,123,212-214,232,239 while one study was performed in a general gynecology practice,231 two studies were performed in a hospital setting,246,247 and the remaining 42 studies were performed at a subspecialty practice.

In addition to the above studies, 6 systematic reviews; 5 good70,84,87,92,95,1 fair quality94 addressed the comparative effectiveness of various treatments in women with unexplained infertility are discussed below and the consistency of their findings with our included studies are incorporated in to our strength of evidence ratings.

Key Points for Unexplained Infertility

Key findings for couples with unexplained infertility included:

- There is no difference between the oral agents of letrozole and anastrozole for the outcome of ectopic pregnancy (low SOE) but evidence is insufficient for other outcomes of interest.
- There is no difference between letrozole and clomiphene for outcomes of multiple births or miscarriage (moderate SOE).
- There is no difference between differing adjunct treatments used in combination with oral agents and IUI for the outcomes of live birth, miscarriage, and OHSS (low SOE for all outcomes).
• There are no differences between immediate IVF versus other treatments prior to IVF for the outcomes of live birth, multiple births, ectopic pregnancy, miscarriage, low birthweight, and OHSS (low SOE for all outcomes). There is however shorter time to pregnancy with immediate IVF (moderate SOE).
• As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE).

**Detailed Synthesis by Treatment for Unexplained Infertility**

Included studies and their findings for the following treatments for unexplained infertility are detailed in this section:

1. Oral Agents Without IUI
2. Oral Agents Versus Unstimulated IUI Versus Expectant Management
3. Different Adjunct Treatments Combined With Oral Agents and IUI
4. Oral Agents With IUI Versus Gonadotropins With IUI
5. Different Treatment Strategies for Controlled Ovarian Hyperstimulation with Gonadotropins and IUI
6. Immediate IVF Versus Other Treatments Prior to IVF
7. Expectant Management Versus Other Interventions
8. ART
   a. IVF
   b. ICSI
   c. IVF Versus ICSI
   d. Unspecified ART

**1. Oral Agents Without IUI for Unexplained Infertility**

Three RCTs, provided evidence on outcomes among women with unexplained infertility treated with oral agents alone. One fair-quality study\(^\text{206}\) compared clomiphene citrate alone to clomiphene citrate plus hydrotubation, another fair-quality study\(^\text{222}\) compared clomiphene to letrozole and anastrozole, a third good-quality trial compared use of letrozole compare with clomiphene with estradiol in women with unexplained fertility who did not respond initially to clomiphene alone.\(^\text{246}\) Results are summarized in Table 18. The studies did not identify significant differences in outcomes between clomiphene alone and comparators, but sample sizes were small, with wide confidence intervals for rates of all outcomes. Outcomes of multiple births, costs, short term adverse effects, and long term outcomes were not evaluated in the included studies.

A good-quality systematic review and meta-analysis compared outcomes between clomiphene citrate and letrozole.\(^\text{70}\) Six studies were included representing 1776 patients, of which one\(^\text{222}\) was included in the present systematic review. The meta-analysis did not assess live birth rates. Miscarriage however was examined in 5 studies (395 patients), including the study by Badawy and colleagues.\(^\text{222}\) Within these studies, the miscarriage rate was 24/195 (12.3%) in the letrozole group versus 36/200 (18.0%) in the clomiphene citrate group (RR 0.65; 95% CI, 0.33 to 1.29, \(p=0.22\)) (moderate SOE). Multiple pregnancy rates were also evaluated in these same 5 trials and again demonstrated no difference between letrozole and clomiphene citrate (RR 0.57, 95% CI 0.25 to 1.27) (moderate SOE).\(^\text{70}\)
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Yapca, 2015</td>
<td>CC</td>
<td>CC + hydrotubation</td>
<td>2/40 (5) (0.6 to 13.5)</td>
<td>8/40 (20) (9.3 to 33.5)</td>
<td>0.043</td>
<td>Greater live births with clomiphene and hydrotubation compared to clomiphene alone</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Badawy, 2009</td>
<td>CC</td>
<td>Letrozole</td>
<td>1/420 (0.2) (0.06 to 0.9)</td>
<td>0/269 (0) (0 to 0.9)</td>
<td>0.43</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Badawy, 2009</td>
<td>CC</td>
<td>Anastrozole</td>
<td>0/107 (0) (0 to 2.3)</td>
<td>0/269 (0) (0 to 0.9)</td>
<td>0.62</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Harira 2018</td>
<td>CC + estradiol</td>
<td>Letrozole</td>
<td>0/86 (0)</td>
<td>0/86 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Yapca, 2015</td>
<td>CC</td>
<td>CC + hydrotubation</td>
<td>1/40 (2.5) (0.06 to 9.0)</td>
<td>0/40 (0) (0 to 6.1)</td>
<td>0.31</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Harira 2018</td>
<td>CC + estradiol</td>
<td>Letrozole</td>
<td>4/86 (4.6)</td>
<td>3/86 (3.4)</td>
<td>0.67</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS</td>
<td>Harira 2018</td>
<td>CC + estradiol</td>
<td>Letrozole</td>
<td>0/86</td>
<td>0/86</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight (small for gestational age)</td>
<td>Badawy, 2009</td>
<td>CC</td>
<td>Letrozole</td>
<td>3/65 (4.6) (1.0 to 10.8)</td>
<td>2/30 (6.7) (0.8 to 17.8)</td>
<td>0.68</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Badawy, 2009</td>
<td>CC</td>
<td>Anastrozole</td>
<td>1/11 (9.1) (95 0.3 to 30.8)</td>
<td>2/30 (6.7) (0.8 to 17.8)</td>
<td>0.54</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal death</td>
<td>Badawy, 2009</td>
<td>CC</td>
<td>Letrozole</td>
<td>0/65 (0) (0 to 3.8)</td>
<td>1/30 (3.3) (0.09 to 11.9)</td>
<td>0.14</td>
<td>No difference</td>
</tr>
</tbody>
</table>
### 2. Oral Agents Versus Unstimulated IUI Versus Expectant Management for Unexplained Infertility

One good-quality RCT\(^{135}\) compared outcomes between oral agents, unstimulated IUI, and expectant management. A second good-quality RCT compared IUI with either clomiphene citrate or letrozole to expectant management.\(^{238}\) Results of both studies are summarized in Table 19 for live birth, pregnancy complications, and time to pregnancy. Other outcomes of interest were not reported.

**Table 19. Outcomes for comparisons of oral agents versus unstimulated IUI versus expectant management in women with unexplained infertility**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhattacharya, 2008(^{135})</td>
<td>RCT (576)</td>
<td>CC</td>
<td>IUI</td>
<td>26/192 (13.5) (9.1 to 18.7)</td>
<td>43/191 (22.5) (16.9 to 28.7)</td>
<td>0.022</td>
<td>Greater live births with IUI compared to clomiphene expectant management</td>
</tr>
<tr>
<td>Farquhar, 2018(^{238})</td>
<td>RCT (201)</td>
<td>IUI CC or letrozole</td>
<td>Expectant management</td>
<td>31/101 (31)</td>
<td>9/100 (9)</td>
<td>0.0001</td>
<td>Greater live births with IUI and oral agents than with expectant management</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Bhattacharya, 2008(^{135})</td>
<td>RCT (576)</td>
<td>CC</td>
<td>0/192 (0) (0 to 1.3)</td>
<td>2/191 (1.0) (0.1 to 2.9)</td>
<td>0.24</td>
<td>No difference</td>
</tr>
<tr>
<td>Farquhar, 2018(^{238})</td>
<td>RCT (201)</td>
<td>IUI CC or letrozole</td>
<td>Expectant management</td>
<td>4/101 (3.9)</td>
<td>0/100 (0)</td>
<td>0.0974</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; N=number of patients; NR=not reported; NS=not statistically significant; OHSS=ovarian hyperstimulation syndrome; RCT=randomized control trial.
Pregnancy complications: Miscarriage

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy complications:</td>
<td>Bhattacharya, 2008</td>
<td>CC</td>
<td>IUI</td>
<td>10/38 (26.3) (13.8 to 41.2)</td>
<td>9/55 (16.3) (7.9 to 27.1)</td>
<td>0.24</td>
<td>No difference</td>
</tr>
<tr>
<td>Miscarriage</td>
<td>RCT (576)</td>
<td></td>
<td>Expectant management</td>
<td>14/46 (30.4) (18.2 to 44.3)</td>
<td></td>
<td>0.68</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Farquhar, 2018</td>
<td>IUI</td>
<td>CC or letrozole</td>
<td>6/37 (16.2)</td>
<td>1/11 (9.1)</td>
<td>0.153</td>
<td>No difference</td>
</tr>
<tr>
<td>Time to pregnancy</td>
<td>Bhattacharya, 2008</td>
<td>CC</td>
<td>IUI</td>
<td>HR of CC compared to expectant managemen 0.83 (0.42 to 1.63)</td>
<td>HR of IUI compared to expectant managemen 1.40 (0.77 to 2.56)</td>
<td>NS</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; HR=hazard ratio; IUI=intrauterine insemination; N=number of patients; NS=not statistically significant; RCT=randomized control trial

This comparison was also evaluated in two good-quality systematic reviews.87,92 Both of these systematic reviews included our included study by Bhattacharya and colleagues.135

One systematic review examined outcomes with clomiphene citrate for unexplained infertility as compared to expectant management.92 Seven RCTs were included (1,159 patients), of which only the one study by Bhattacharya135 reported on outcomes of interest. In this study, there was no benefit of clomiphene citrate over placebo for live birth (OR 0.79; 95% CI, 0.44 to 1.38, p=0.41) or miscarriage (OR = 0.71, 95% CI 0.31 to 1.61). The SOE was rated as insufficient given imprecise findings from one study.

Another good-quality systematic review of 14 studies examined outcomes following ovarian stimulation, IUI, or both in 2,033 patients with unexplained infertility.87 Comparisons were made between IUI and timed intercourse (TI) with and without ovarian hyperstimulation. Clomiphene citrate, gonadotropins, or a combination of clomiphene citrate and gonadotropins were utilized for ovarian hyperstimulation. Fourteen RCTs were included, of which one135 was included in our systematic review. Overall for the 14 included studies, they noted that the risk of bias was “substantial” due to failure to report allocation concealment and details of randomization.

When comparing IUI versus expectant management, only one good-quality, low risk of bias study was included, also included in the present systematic review.135 Live birth rates were not significantly differently between groups (23% with IUI, 16% with expectant management; OR 1.60; 95% CI, 0.92 to 2.78). Miscarriage rates were similar between groups (OR 0.77; 95% CI, 0.28 to 2.11). There were 2 ectopic pregnancies in the IUI group (OR 5.06; 95% CI, 0.24 to 106.21) (insufficient SOE).

For the comparison of IUI versus TI with ovarian hyperstimulation, only 2 studies examined live birth rates. Live birth rates were similar between groups (OR 1.59; 95% CI, 0.88 to 2.88).

For the comparison of IUI in a natural cycle versus IUI in a stimulated cycle, 3 studies examined live birth rates, all excluded from the present review due to publication prior to 2007.
Live birth rates were significantly higher with ovarian hyperstimulation/IUI compared to IUI alone (OR 2.07; 95% CI, 1.22 to 3.50) (low SOE).

Finally, only one study was included for in the meta analysis which compared IUI in a natural cycle versus TI in a stimulated cycle (also included in the present study).135 Odds of live birth were higher with IUI than with ovarian hyperstimulation/TI (OR 1.95; 95% CI, 1.10 to 3.44) (insufficient SOE).

Overall this systematic review suggested that there is evidence that IUI with ovarian hyperstimulation increases the live birth rate compared to IUI alone. Other comparisons had insufficient SOE.

3. Different Adjunct Treatments Combined With Oral Agents or Gonadotropins and IUI for Unexplained infertility

Two good-quality RCTs227,229 and six fair-quality RCTs216,221,231,232,242,245 compared outcomes between different adjunct treatment strategies for use of oral agents (usually CC) and/or gonadotropins with IUI. Interventions included luteal progesterone using a vaginal pessary,229 hydrotubation with saline compared to lidocaine,227 double versus single insemination,221, piroxicam231, vaginal progesterone versus oral dygesterone232, supine immobilization,242 cervical mucus removal,245 and adjunct stimulation with either rFSH or HMG.216 Results are summarized in Table 20. Other than an increase in live birth for women in the cervical mucus removal group (RR 1.60, 95% CI 1.12 to 2.28), none of the studies identified significant differences in outcomes between interventions for the outcomes of live birth, miscarriage, or OHSS (low SOE for all outcomes). No other outcomes of interest were evaluated.

Table 20. Outcomes for comparisons of oral agents with IUI in women with unexplained infertility

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>p Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Ebrahimi, 2010229 RCT (200)</td>
<td>CC/hMG + Progesterone</td>
<td>CC/hMG + No support</td>
<td>19/98 (19.4) (12.2 to 27.7)</td>
<td>15/102 (14.7) (8.6 to 22.2)</td>
<td>0.38</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Bagis, 2010231 RCT (228)</td>
<td>CC+Double insemination</td>
<td>CC+Single insemination</td>
<td>1/19 (5.3) (0.1 to 18.5)</td>
<td>2/17 (11.7) (1.5 to 30.2)</td>
<td>0.48</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Rashidi, 2013216 RCT (259)</td>
<td>CC+rFSH</td>
<td>CC+hMG</td>
<td>19/132 (14.5) (9.0 to 21.0)</td>
<td>16/127 (12.6) (7.4 to 18.9)</td>
<td>0.12</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>van Rijswijk, 2017242 RCT (481)</td>
<td>Supine immobilizatio n + IUI</td>
<td>Immediate mobilization + IUI</td>
<td>73/226 (32)</td>
<td>92/245 (37)</td>
<td>0.13</td>
<td>No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Morad, 2012^{227} RCT (216)</td>
<td>Lidocaine hydrotubation prior to CC + IUI</td>
<td>Saline hydrotubation prior to CC + IUI</td>
<td>1/109 (0.9) (0.02 to 3.4)</td>
<td>1/107 (0.9) (0.02 to 3.4)</td>
<td>0.99</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Bagis, 2010^{231} RCT (228)</td>
<td>CC+Double insemination</td>
<td>CC+Single insemination</td>
<td>6/20 (30) (12.6 to 51)</td>
<td>4/16 (25) (7.8 to 48.1)</td>
<td>0.74</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Morad, 2012^{227} RCT (216)</td>
<td>Lidocaine hydrotubation prior to CC + IUI</td>
<td>Saline hydrotubation prior to CC + IUI</td>
<td>1/109 (0.9) (0.02 to 3.4)</td>
<td>1/107 (0.9) (0.02 to 3.4)</td>
<td>0.99</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Zarei, 2016^{231} RCT (260)</td>
<td>Piroxicam+ CC + rFSH + hCG + IUI</td>
<td>Placebo+ CC + rFSH + hCG + IUI</td>
<td>5/130 (4)</td>
<td>5/130 (4)</td>
<td>0.82</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Khosravi, 2015^{232} RCT (150)</td>
<td>Oral dygesterone + IUI</td>
<td>Vaginal progesterone + IUI</td>
<td>2/75 (9.1)</td>
<td>3/75 (15.8)</td>
<td>0.056</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Maher, 2018^{243} RCT (714)</td>
<td>Cervical mucus removal</td>
<td>No mucus removal</td>
<td>14/361 (12.5)</td>
<td>11/353 (14.3)</td>
<td>0.72</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS</td>
<td>Morad, 2012^{227} RCT (216)</td>
<td>Lidocaine hydrotubation prior to CC + IUI</td>
<td>Saline hydrotubation prior to CC + IUI</td>
<td>1/109 (0.9) (0.02 to 3.4)</td>
<td>1/107 (0.9) (0.02 to 3.4)</td>
<td>0.99</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Rashidi, 2013^{216} RCT (259)</td>
<td>CC+rFSH</td>
<td>CC+hMG</td>
<td>6/132 (4.5) (1.7 to 8.7)</td>
<td>5/127 (3.9) (1.3 to 7.9)</td>
<td>0.81</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Maher, 2018^{243} RCT (714)</td>
<td>Cervical mucus removal</td>
<td>No mucus removal</td>
<td>18/361 (5)</td>
<td>15/353 (4.2)</td>
<td>0.61</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; hMG=human menopausal gonadotropin; IUI=intrauterine insemination; N=number of patients; NR=not reported; NS=not statistically significant; OHSS=ovarian hyperstimulation syndrome; rFSH=recombinant follicle-stimulating hormone; RCT=randomized control trial
4. Oral Agents With IUI Versus Gonadotropins With IUI for Unexplained Infertility

Six RCTs (4 good, one fair, and one poor quality) compared outcomes between oral agents/IUI and gonadotropins/IUI. Table 21 summarizes the findings from these studies. Evidence supported no difference in miscarriage rates between strategies (low SOE). All other outcomes had inconsistent and imprecise findings resulting in insufficient SOE.

Table 21. Outcomes for comparisons of oral agents with IUI versus gonadotropins with IUI for unexplained infertility

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%)</th>
<th>Results Comparator N (%)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Erdem, 2015 (174)</td>
<td>CC</td>
<td>rFSH with IUI</td>
<td>42/94 (44.7)</td>
<td>25/87 (28.7)</td>
<td>0.026</td>
<td>Greater live births with CC</td>
</tr>
<tr>
<td>Gregoriou, 2008 (50)</td>
<td>Letrozole</td>
<td>rFSH with IUI</td>
<td>7/25 (28)</td>
<td>5/25 (20)</td>
<td>0.51</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Diamond, 2015 (900)</td>
<td>CC with IUI</td>
<td>Subcutaneous gonadotropin with IUI</td>
<td>70/300 (23.3)</td>
<td>97/301 (32.2)</td>
<td>0.02</td>
<td>Greater live birth with gonadotropins compared to CC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Letrozole with IUI</td>
<td></td>
<td>56/299 (18.7)</td>
<td></td>
<td>0.000</td>
<td>Greater live birth with gonadotropins compared to letrozole</td>
<td></td>
</tr>
<tr>
<td>Danhof, 2018 (738)</td>
<td>CC with IUI</td>
<td>FSH with IUI</td>
<td>92/369 (25)</td>
<td>105/369 (28)</td>
<td>0.36</td>
<td>No significant difference</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Pourali, 2017 (180)</td>
<td>CC + HMG + IUI</td>
<td>Letrozole + hMG + IUI</td>
<td>5/87 (5.7)</td>
<td>4/83 (4.8)</td>
<td>0.80</td>
<td>No difference</td>
</tr>
<tr>
<td>Diamond, 2015 (900)</td>
<td>CC with IUI</td>
<td>Subcutaneous gonadotropin with IUI</td>
<td>31/106 (29.3)</td>
<td>51/140 (36.4)</td>
<td>0.24</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Letrozole with IUI</td>
<td></td>
<td>26/85 (30.6)</td>
<td></td>
<td>0.37</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Danhof, 2018 (738)</td>
<td>CC with IUI</td>
<td>FSH with IUI</td>
<td>31/369 (8)</td>
<td>32/369 (9)</td>
<td>0.63</td>
<td>No difference</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%)</td>
<td>Results Comparator N (%)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>---------------------------</td>
<td>-----------------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>---------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Pregnancy complications:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple births</td>
<td>Diamond, 2015(^{233})</td>
<td>CC with IUI</td>
<td>Subcutaneous gonadotropin with IUI</td>
<td>4/70 (5.7)</td>
<td>31/97 (32.0)</td>
<td>&lt;0.00 01</td>
<td>Greater multiple gestations with gonadotropins compared to CC</td>
</tr>
<tr>
<td></td>
<td>RCT (900)</td>
<td></td>
<td>Letrozole with IUI</td>
<td>8/56 (14.3)</td>
<td></td>
<td>0.015</td>
<td>Greater multiple gestations with gonadotropins compared to CC</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term adverse effects:</td>
<td>Nada, 2016(^{234})</td>
<td>CC with IUI</td>
<td>GnRH antagonist with IUI</td>
<td>6/297 (2)</td>
<td>30/298 (10)</td>
<td>&lt;0.00 01</td>
<td>Higher rate of mild OHSS among those in GnRH antagonist group</td>
</tr>
<tr>
<td>OHSS</td>
<td>RCT (595)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pourali, 2017(^{241})</td>
<td>CC + HMG + IUI</td>
<td>Letrozole + hMG + IUI</td>
<td>5/87 (5.7)</td>
<td>0/83 (0)</td>
<td>0.027</td>
<td>Rate of cancelled cycles for OHSS higher in CC group</td>
</tr>
<tr>
<td></td>
<td>RCT (170)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ectopic pregnancy</td>
<td>Danhof, 2018(^{244})</td>
<td>CC with IUI</td>
<td>FSH with IUI</td>
<td>3/369 (1)</td>
<td>2/369 (1)</td>
<td>1</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>RCT (738)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CC=clomiphene citrate; CI=confidence interval; FSH=follicle-stimulating hormone; hMG=human menopausal gonadotropin; IUI=intrauterine insemination; N=number of patients; NS=not statistically significant; OHSS=ovarian hyperstimulation syndrome; RCT=randomized control trial; rFSH=recombinant follicle-stimulating hormone; uFSH=urinary follicle-stimulating hormone

5. Different Treatment Strategies for Controlled Ovarian Hyperstimulation With Gonadotropins and IUI for Unexplained Infertility

Two good-quality RCTs,\(^{209,230}\) two fair-quality RCTs\(^{221,225}\) and one poor-quality RCT\(^{211}\) compared outcomes between different treatment strategies of gonadotropins/IUI. Interventions included luteal progesterone compared to no luteal support,\(^{209,225}\) uterine cavity perturbation versus no additional therapy,\(^{211}\) and type of gonadotropin (rFSH, highly purified urinary FSH, or HMG).\(^{230}\) Results are summarized in Table 22. The studies were consistent in terms of the direction of benefit (greater live births with progesterone compared to no support) but inconsistent in terms of their statistical significance and with imprecise findings (insufficient SOE). No difference was reported between uterine perturbation versus no intervention.
(insufficient SOE). The imprecise findings from heterogeneous interventions for miscarriage resulted in an insufficient strength of evidence rating. Finally, evidence from one study with no events provided insufficient evidence to support statements about the impact on OHSS. No other outcomes of interest were reported.

**Table 22. Outcomes for comparisons of different treatment strategies for controlled ovarian hyperstimulation with gonadotropins and IUI in women with unexplained infertility**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Seckin, 2014209 (149)</td>
<td>Progesterone</td>
<td>No support</td>
<td>14/71 (19.7)</td>
<td>11/78 (14.1)</td>
<td>0.36</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Erdem, 2009223 (214)</td>
<td>Progesterone</td>
<td>No support</td>
<td>39/223 (per cycle) (17.5)</td>
<td>19/204 (per cycle) (9.3)</td>
<td>0.016</td>
<td>Greater live births with progesterone compared to no support</td>
</tr>
<tr>
<td></td>
<td>Yildiz, 2014211 (180)</td>
<td>Uterine perturbation</td>
<td>No intervention</td>
<td>10/79 (12.7) (6.3 to 20.8)</td>
<td>20/101 (19.8) (12.6 to 28.1)</td>
<td>0.20</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Bagis, 2010221 (228)</td>
<td>Double insemination</td>
<td>Single insemination</td>
<td>6/20 (30) (12.6 to 51)</td>
<td>4/16 (25) (7.8 to 48.1)</td>
<td>0.74</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Yildiz, 2014211 (180)</td>
<td>Uterine perturbation</td>
<td>No intervention</td>
<td>1/79 (1.3) (0.03 to 4.5)</td>
<td>3/101 (3.0) (0.6 to 7.0)</td>
<td>0.44</td>
<td>No difference</td>
</tr>
<tr>
<td>Short-term adverse effects: OHSS</td>
<td>Demirol, 2007230 (241)</td>
<td>rFSH</td>
<td>hMG</td>
<td>2/81 (2.5) (0.3 to 6.8)</td>
<td>hMG: 2/80 (2.5) (0.3 to 6.9)</td>
<td>0.99</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>uFSH</td>
<td>uFHS: 1/80 (1.2) (0.03 to 4.5)</td>
<td>uFHS: 0/80 (0) (0 to 3.1)</td>
<td>0.57</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NS</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; FSH=follicle-stimulating hormone; hMG=human menopausal gonadotropin; IUI=intrauterine insemination; N=number of patients; NS=not statistically significant; OHSS=ovarian hyperstimulation
syndrome; RCT=randomized control trial; rFSH=recombinant follicle-stimulating hormone; uFSH=urinary follicle-stimulating hormone

6. Immediate IVF Versus Other Treatments Prior to IVF for Unexplained Infertility

Three RCTs (2 good\textsuperscript{52,210} 1 fair quality\textsuperscript{240}) compared different broad strategies for women with unexplained infertility. In the Fast Track and Standard Treatment (FASTT) trial,\textsuperscript{52} women aged 21-39 were randomized to (a) up to 3 cycles of clomiphene citrate/IUI followed by up to 3 cycles of gonadotropins/IUI followed by up to 6 cycles of IVF (“conventional strategy”) or (b) up to 3 cycles of clomiphene citrate/IUI followed by 6 cycles of IVF (“fast track”). In the Forty and Over Treatment Trial (FORT-T),\textsuperscript{210} women aged 38-42 with unexplained infertility were randomized to (a) clomiphene citrate/IUI for up to 2 cycles followed by up to 6 cycles IVF, (b) FSH/IUI for up to 2 cycles followed by 6 cycles IVF, or (c) immediate IVF (up to 6 cycles). A final fair-quality RCT randomized 207 couples with unexplained infertility to 3 cycles of IUI plus controlled ovarian hyperstimulation with injectable gonadotropins or to 1 cycle of IVF.\textsuperscript{240}

Table 23 presents outcomes of interest. Consistent findings between the three studies found no difference in outcomes between immediate IVF compared with other treatments prior to IVF (low SOE for live birth, ectopic, miscarriage, multiple births, low birthweight, OHSS) while decreasing the time to pregnancy (moderate SOE).

\begin{table}[h]
\centering
\begin{tabular}{|l|l|l|l|l|l|l|}
\hline
\textbf{Outcome} & \textbf{Study Design (N Patients)} & \textbf{Intervention} & \textbf{Comparator} & \textbf{Results Intervention N (95\% CI)} & \textbf{Results Comparator N (95\% CI)} & \textbf{P Value} & \textbf{Summary of Study Findings} \\
\hline
Live birth—by strategy & Reindollar, 2010\textsuperscript{52} RCT (503) & CC\rightarrow IVF (fast track) & CC\rightarrow gonadotropins \rightarrow IVF (conventional) & 171/256 (66.8) (60.5 to 72.1) & 150/247 (60.7) (56.6 to 68.9) & 0.15 & No difference \\
Goldman, 2014\textsuperscript{210} RCT (154) & IVF & CC/IUI\rightarrow IVF & Gonadotropins / IUI\rightarrow IVF & 24/51 (47.1) (31.2 to 63.4) & 25/51 (49.0) (32.9 to 65.2) & 0.84 & No difference \\
Nandi, 2017\textsuperscript{240} RCT (207) & 3 cycles IUI + ovarian hyperstimulation with gonadotropins & IVF & & 29/101 (28.7) & 36/106 (33.9) & 0.42 & No difference \\
\hline
Pregnancy complications: Ectopic pregnancy—by strategy & Reindollar, 2010\textsuperscript{52} RCT (503) & CC\rightarrow IVF (fast track) & CC\rightarrow gonadotropins \rightarrow IVF (conventional) & 10/256 (3.9) (1.9 to 6.6) & 8/247 (3.2) (1.4 to 5.8) & 0.69 & No difference \\
\hline
\end{tabular}
\caption{Outcomes for comparisons of immediate IVF versus other treatments prior to IVF in women with unexplained infertility\textsuperscript{a}}
\end{table}
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P Value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Goldman, 2014 (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>0/51 (0) (95 CI 0 to 4.8)</td>
<td>CC/IUI→IVF: 1/51 (2.0) (0.04 to 7.1)</td>
<td>0.32</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF: 3/52 (5.8) (1.2 to 13.5)</td>
<td>0.08</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Nandi, 2017 (207)</td>
<td>3 cycles IUI + ovarian hyperstimulation with gonadotropins</td>
<td>IVF</td>
<td>2/101 (1.98)</td>
<td>0/106 (0)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Reindollar, 2010 (503)</td>
<td>CC→IVF (fast track)</td>
<td>CC→gonadotropins → IVF (conventional)</td>
<td>38/256 (14.8) (10.8 to 19.4)</td>
<td>32/247 (13.0) (9.1 to 17.4)</td>
<td>0.54</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Goldman, 2014 (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>11/51 (21.6) (11.5 to 33.7)</td>
<td>CC/IUI→IVF: 8/51 (15.7) (7.2 to 26.7)</td>
<td>0.45</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nandi, 2017 (207)</td>
<td>3 cycles IUI + ovarian hyperstimulation with gonadotropins</td>
<td>IVF</td>
<td>3/34 (12)</td>
<td>13/49 (26.5)</td>
<td>0.11</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Low birthweight</td>
<td>Reindollar, 2010 (503)</td>
<td>CC→IVF (fast track)</td>
<td>CC→gonadotropins → IVF (conventional)</td>
<td>34/171 (19.9) (14.3 to 26.2)</td>
<td>30/150 (twins) (20) (14.0 to 26.7)</td>
<td>0.98</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Goldman, 2014 (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>3/24 (twins) (2.8 to 28.0)</td>
<td>CC/IUI→IVF: 6/25 (twins) (25) (9.8 to 42.2)</td>
<td>0.30</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF: 4/22 (3 twins, 1 triplet) (18.2) (5.4 to 36.3)</td>
<td>0.59</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reindollar, 2010 (503)</td>
<td>CC→IVF (fast track)</td>
<td>CC→gonadotropins → IVF (conventional)</td>
<td>30/171 (17.5) (95 CI 12.2 to 23.6)</td>
<td>23/150 (15.3) (10.0 to 21.5)</td>
<td>0.59</td>
<td>No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P Value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------------------------</td>
<td>--------------</td>
<td>-----------------------------</td>
<td>-------------------------------------</td>
<td>-----------------------------------</td>
<td>---------</td>
<td>---------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Goldman, 2014&lt;sup&gt;218&lt;/sup&gt;</td>
<td>RCT (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>2/24 (8.3) (1.1 to 21.9)</td>
<td>CC/IUI→IVF: 2/25 (2.8) (1.0 to 21.1)</td>
<td>0.97</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gonadotropins / IUI→IVF</td>
<td>Gonadotropins / IUI→IVF: 2/22 (9.1) (1.2 to 23.8)</td>
<td>0.93</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Neonatal deaths</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reindollar, 2010&lt;sup&gt;52&lt;/sup&gt;</td>
<td>RCT (503)</td>
<td>CC→IVF (fast track)</td>
<td>CC→gonadotropins → IVF (conventional)</td>
<td>0/171 (0) (0 to 1.5)</td>
<td>0/150 (0) (0 to 1.7)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Goldman, 2014&lt;sup&gt;218&lt;/sup&gt;</td>
<td>RCT (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>0/24 (0) (0 to 10.0)</td>
<td>CC/IUI→IVF: 0/25 (0) (0 to 9.6)</td>
<td>NS</td>
<td>No difference</td>
</tr>
<tr>
<td>Time to pregnancy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reindollar, 2010&lt;sup&gt;52&lt;/sup&gt;</td>
<td>RCT (503)</td>
<td>CC→IVF (fast track)</td>
<td>CC→gonadotropins → IVF (conventional)</td>
<td>8 months</td>
<td>11 months</td>
<td>0.045</td>
<td>Shorter time to pregnancy with immediate IVF compared to gonadotropins prior to IVF</td>
</tr>
<tr>
<td>Goldman, 2014&lt;sup&gt;218&lt;/sup&gt;</td>
<td>RCT (154)</td>
<td>IVF</td>
<td>CC/IUI→IVF</td>
<td>8.7 ± 0.5 months</td>
<td>CC/IUI→IVF: 9.1 ± 0.6</td>
<td>&lt;0.05</td>
<td>Shorter time to pregnancy with immediate IVF compared to IUI and gonadotropins prior to IVF</td>
</tr>
</tbody>
</table>
## 7. Expectant Management Versus Other Interventions for Unexplained Infertility

We identified three studies that included expectant management as a strategy and provided evidence on at least one outcome of interest.

One good-quality RCT examined outcomes following luteal phase scratching compared to expectant management in women with unexplained infertility. The miscarriage rate was 2/16 (12.5%) in the endometrial scratch group versus 1/6 (16.6%) in the control group (p=0.79). Multiple births were also similar in the two groups (1/54 in the endometrial scratch group versus 0/51 in the control group, p=1) (insufficient SOE for both outcomes).

Another good-quality RCT trial of 253 couples compared costs of controlled ovarian stimulation with FSH and IUI to 6 months of expectant management followed by usual treatment. Three-year outcomes were evaluated. The mean estimated costs per couple in the expectant management group were 3424 euros (95% CI, 880 to 5968 euros) and in the immediate treatment group 6040 euros (95% CI, 4055-8121 euros), with an estimated saving of 2616 euros per couple (95% CI, 385 to 4847 euros). The likelihood of achieving pregnancy and the time to pregnancy did not differ between the groups, suggesting that expectant management was a reasonable cost-savings option. Applicability to U.S. populations is limited because the study was conducted in the Netherlands (insufficient SOE).

A third study was a good-quality RCT from the Netherlands which randomized women with different types of infertility, including unexplained infertility, to a 6-month structured lifestyle intervention for weight loss or to prompt treatment for infertility as per Dutch infertility guidelines.
Of the 161 women with unexplained infertility, 86 were assigned to the intervention and 77 to the control groups. Within this group, there were no significant differences in the outcomes of vaginal birth of healthy singleton at term or of live births. Compared to the control group, the RR (95% CI) of vaginal birth of health singleton at term for the intervention group was 0.73 (0.47 to 1.1); compared to the control group, the RR (95% CI) of live births for the intervention group was 0.85 (0.64 to 1.1).119

Finally, a good-quality systematic review84 examined outcomes following IVF for unexplained infertility. Within this analysis, the live birth rate was higher with IVF than expectant management (45.8% vs. 3.7%, OR 22.00; 95% CI, 2.56 to 189.37) based on 1 RCT with 51 women. Given the inconsistent findings we rated the SOE as insufficient.

8. ART for Unexplained Infertility

IVF for Unexplained Infertility

Two observational studies236,239 and one RCT247 evaluated different IVF strategies in women with unexplained infertility.

The fair quality RCT compared 3 strategies for ovarian stimulation in women with diminished ovarian reserve undergoing IVF or ICSI.247 There were 116 participants randomized to one of the 3 protocols: a modified GnRH agonist (triptorelin) protocol; a mild stimulation protocol with letrozole; or an antagonist protocol with triptorelin. There was no significant difference in live birth rate between the 3 groups, with rates of 20.37% (11/54), 15.38% (8/52) and 13.33% (8/60) respectively.247

One good-quality observational study utilizing linked data from SMART Collaborative examined rates of small for gestational age with fresh embryo transfer, frozen embryo transfer, and natural conception.239 Among women with unexplained infertility, IVF with fresh embryo transfer was associated with increased odds of small for gestational age (defined as <10%) compared to natural conception (adjusted OR 1.24, 95% CI 1.10 to 1.38). Results were similar when small for gestational age was defined as <5% (adjusted OR 1.24, 95% CI 1.06 to 1.45). When IVF was performed with frozen embryo transfer, however, there was no significant difference in small for gestational age compared to natural conception (adjusted OR 0.71, 95% CI 0.47 to 1.06 when defined as <10%; adjusted OR 0.78, 95% CI 0.45 to 1.36 when defined as <5%).

The second good-quality observational study evaluated low birth weight and demonstrated that among women with unexplained infertility, infants conceived with ART had an increased risk of low birth weight (OR 1.45, 95% CI 1.29 to 1.63).236 However, a discordant sibling-pair analysis in the same study demonstrated no association (OR 1.27, 95% CI 0.95 to 1.68)

A fair-quality systematic review and meta-analysis examined outcomes following ART with and without preimplantation genetic screening (PGS) in women of advanced maternal age (35+).94 Six RCTs were included (1,136 patients), none of the studies within this systematic review were included in the present systematic review. Reasons for exclusion included either being published before 2007, not presenting outcomes of interest by underlying diagnosis, or presenting findings by embryo rather than by cycle or patient. Live birth rate per woman was examined in only one study. In this study the odds of live birth were decreased with ART plus PGS compared with ART without PGS (OR 0.48; 95% CI, 0.26 to 0.88). Miscarriage was examined in 3 studies. There was no significant difference in miscarriage rate between groups (OR 2.02; 95% CI, 0.57 to 7.41). The systematic review was limited by inclusion of studies that
utilized PGS technologies that are no longer current. Therefore, the review is not applicable to today’s clinical practice and strength of evidence for all outcomes was rated as insufficient.

**ICSI for Unexplained Infertility**

One fair-quality RCT\(^2\) compared outcomes between different treatment strategies with ICSI. The study randomized patients undergoing ICSI to hyaluronic acid sperm selection (physiological ICSI [PICSI]) or no selection (ICSI).\(^2\) The live birth rate per patient was 22/71 (31%) with PICSI versus 21/80 (26.3%) with ICSI (p=0.520) (insufficient SOE). The miscarriage rate per patient was 3/25 (12%) with PICSI versus 7/28 (25%) with ICSI (p=0.227) (insufficient SOE).

**IVF Versus Combined IVF and ICSI for Unexplained Infertility**

Two observational studies evaluated outcomes for women undergoing IVF as compared to combined IVF and ICSI for unexplained infertility.\(^2\)

In one good-quality observational study\(^2\) the authors performed an adaptive decision analysis to determine whether combined IVF and split-ICSI (randomly assigning sibling oocytes to conventional IVF or combined IVF and ICSI) is cost-effective for patients with unexplained infertility undergoing their first IVF cycle. For patients undergoing one cycle, conventional IVF was preferred, as the incremental cost effectiveness ratio (ICER) of split-ICSI or all ICSI ($58,766 per additional live birth) did not justify the increase in live birth rate (3%). For patients undergoing two cycles, split IVF/ICSI was preferred, as the 3.3% increased cumulative live birth rate was gained at an ICER of $29,666 per additional live birth (insufficient SOE).

The second observational study utilizing data from SART CORS compared rates of preterm birth and low birth weight among women undergoing conventional IVF versus ICSI.\(^2\) In this study, 2,922 live births were to women with unexplained infertility. In a matched dataset, among those women with unexplained infertility, the proportion of births with low birth weight was 123/1,464 (8.4%) in the ICSI group compared to 124/1,458 in the conventional group (8.5%) (OR 1.03, 95% CI 0.73 to 1.47, P=0.86) demonstrating no evidence of a difference between treatment groups (low SOE).

**Unspecified ART for Unexplained Infertility**

**Long-Term Outcomes After ART for Unexplained Infertility: Child**

Risk for cancer in children conceived after by either IVF or ICSI was evaluated in a cohort of 106,013 children born in Britain between 1992 and 2008.\(^2\) For 33,840 of these children, the cause of infertility was unexplained. The average duration of follow-up was 6.6 years, with a maximum of 15 years. Cancer diagnoses were identified through the National Registry of Childhood Tumors.

Cancer was diagnosed in 32 children born after assisted conception for unexplained infertility. The standardized incidence ratio (SIR) for all cancers in children conceived with assisted conception for unexplained infertility compared to the general population of the same age was 0.84 (95% CI, 0.57 to 1.18).

Despite the large number of children in the cohort, the overall number of cancers was small. Conclusions about the risk are tenuous due to the inability to examine risk for individual cancer types by underlying cause of infertility (low SOE).
**Subgroups of Interest for Unexplained Infertility**

Five total studies, one fair-quality observational study, two fair-quality RCTs, and two good-quality RCTs reported data on outcomes with IVF in subgroups of interest for women with infertility of unknown etiology.

**Ovarian Reserve Status**

A fair-quality observational SART CORS study assessed the impact of assisted hatching IVF (where an embryologist uses micromanipulation under a microscope to create a small hole in the zona pellucida) outcomes in patients undergoing an initial IVF cycle with diminished ovarian reserve (DOR). DOR was diagnosed in one of 2 ways: (1) only elevated FSH and (2) DOR diagnosis in SART CORS database. Among women with elevated FSH only, the live birth rate per cycle was 562/2682 (21%) with no assisted hatching compared to 571/3470 (21.6%) with assisted hatching (RR 1.03; 95% CI, 0.94 to 1.14). After adjustment for possible confounders, the RR was 0.92 (95% CI, 0.79 to 1.07). Among women with a DOR diagnosis in SART CORS, the live birth rate per cycle was 3190/14,106 (22.6%) with no assisted hatching compared to 3123/16,033 (19.5%) with assisted hatching (RR 0.86; 95% CI, 0.82 to 0.9). After adjustment for possible confounders, the RR was 0.84 (95% CI, 0.77 to 0.92), suggesting worse outcomes with assisted hatching although the limitations of the observational design remain (insufficient SOE).

A good-quality RCT examined the utility of corifollitropin alpha for ovarian stimulation for poor responders undergoing IVF. Patients were randomized to clomiphene citrate and corifollitropin alpha for the first 7 days of stimulation, followed by recombinant FSH, or to clomiphene citrate with daily FSH. There was no evidence of a difference in live birth rate per transfer was (OR 0.76, 95% CI 0.19 to 3.04, P=0.73) nor in miscarriage rates (OR 1.14, 95% CI 0.06 to 21.87, P=1.00).

A good-quality RCT evaluated women ≥ 35 years old and with a low antral follicular count of < 5 follicles or poor ovarian response or cycle cancellation during a previous IVF irrespective of age and compared the effects of a mild ovarian stimulation protocol compared to a regular ovarian stimulation protocol. There was a similar risk of pregnancy loss in the 2 groups; compared to the conventional strategy, those randomized to the mild stimulation protocol had an RR (95% CI) of early pregnancy loss of 1.20 (0.36 to 4.17).

Finally, one fair-quality RCT examined the effects of melatonin versus placebo on outcomes with IVF in women receiving ART for the first time and with diminished ovarian reserve. Women randomized to the melatonin arm received 3mg at bedtime beginning on day 5 of their cycle prior to the cycle planned for gonadotropin stimulation; women in the placebo group received matching placebo capsules taken similarly. IVF was performed using the same protocol in both groups. Of 32 women in the melatonin arm, 6.2% had a miscarriage; of 34 women in the placebo arm 2.9% had a miscarriage demonstrating no evidence of a difference between treatment groups.

**Advanced Maternal Age**

The study by Rubio and colleagues compared PGS prior to blastocyst transfer on day 3 to no screening in women with advanced maternal age. In the advanced maternal age group, the live birth rate per patient was 30/93 (32.5%) in the PGS group versus 14/90 (15.5%) in the no-PGS group (OR 2.59; 95% CI, 1.26 to 5.30) demonstrating an increase in live birth with PGS (insufficient SOE).
Miscarriage rates and multiple birth rates were similar. In the advanced maternal age population, the miscarriage rate was 16.7% in the PGS group versus 22.2% in the no-PGS group (p>0.05). The twin birth rate was 25% in the PGS group versus 21.4% in the no-PGS group (p>0.05) (insufficient SOE for all outcomes).

Strength of Evidence for Unexplained Infertility

Table 24 summarizes the SOE for the findings described above. SOE for most outcomes was judged to be insufficient or low, primarily because of imprecision or small numbers of studies of fair quality. One exception was time to pregnancy between different strategies for sequencing treatment, where precision was reasonable, and where SOE was judged to be moderate.

Table 24. Strength of evidence for major outcomes—KQ 3 (unexplained infertility)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral Agents Without IUI</td>
<td>Live birth (any/patient)</td>
<td>1 RCT&lt;sup&gt;296&lt;/sup&gt; (80)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 small study with moderate risk of bias.</td>
<td>Insufficient (Imprecise findings, one small study with moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>2 RCTs&lt;sup&gt;223,226&lt;/sup&gt; (1,168)</td>
<td>No difference: No difference between letrozole and anastrozole:</td>
<td>Low (Moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 SR&lt;sup&gt;70&lt;/sup&gt; (5 studies, 395 patients)</td>
<td>No difference: No difference between letrozole and clomiphene citrate</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs&lt;sup&gt;206,223,224,226&lt;/sup&gt; (1,248)</td>
<td>No difference: No difference between letrozole and clomiphene citrate</td>
<td>Moderate</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Death</td>
<td>1 RCT&lt;sup&gt;222&lt;/sup&gt; (996)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations))</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT&lt;sup&gt;222&lt;/sup&gt; (996)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Short term adverse effects of treatment: OHSS</td>
<td>1 RCT&lt;sup&gt;226&lt;/sup&gt; (172)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 small study</td>
<td>Insufficient (Imprecise, one small study)</td>
<td></td>
</tr>
<tr>
<td>Clomiphene Citrate vs. Expectant Management</td>
<td>Live birth</td>
<td>2 RCTs&lt;sup&gt;135,238&lt;/sup&gt; (781)</td>
<td>Inconclusive: SOE was insufficient given inconsistent evidence from studies with heterogeneous interventions.</td>
<td>Insufficient (Inconsistent, heterogeneous interventions)</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic Pregnancy</td>
<td>2 RCTs&lt;sup&gt;135,238&lt;/sup&gt; (781)</td>
<td>No difference: No significant difference in ectopic pregnancy rates between clomiphene and expectant management</td>
<td>Low (Imprecise, heterogeneous interventions)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>2 RCTs&lt;sup&gt;135,238&lt;/sup&gt; (781)</td>
<td>No difference: No significant difference in ectopic pregnancy rates between clomiphene and expectant management</td>
<td>Low (Imprecise, heterogeneous interventions)</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>---------------------------</td>
<td>------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Oral Agents vs. Unstimulated IUI vs. Expectant Management</td>
<td>Live birth</td>
<td>1 SR²⁶³ (3 studies, 370)</td>
<td>Improvement: A significant increase in live births was found for women treated with IUI and ovarian hyperstimulation compared to women treated with IUI only</td>
<td>Low (Inconsistent)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT¹³⁵ (580)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 trial.</td>
<td>Insufficient (Imprecise, one study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT¹³⁵ (580)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 trial.</td>
<td>Insufficient (Imprecise, one study)</td>
</tr>
<tr>
<td></td>
<td>Time to pregnancy</td>
<td>1 RCT¹³⁵ (580)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 trial.</td>
<td>Insufficient (Imprecise, one study)</td>
</tr>
<tr>
<td>Adjunct Treatments with Oral Agents and IUI</td>
<td>Live birth</td>
<td>5 RCTs²¹⁶,²²¹,²²⁹,²⁴²,²⁴⁵ (1859)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT²²⁷ (216)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 trial.</td>
<td>Insufficient (Imprecise, one study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>5 RCTs²²¹,²²⁷,²³¹,²³²,²⁴⁵ (1859)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>3 RCTs²¹⁶,²²⁷,²⁴⁵ (1189)</td>
<td>No difference: No difference between adjunct treatments with oral agents and IUI</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>Oral Agents With IUI vs. Gonadotropins With IUI</td>
<td>Live birth</td>
<td>4 RCTs²⁰⁸,²²⁶,²³³,²⁴⁴ (1708)</td>
<td>Inconclusive: Conflicting findings from RCTs resulted in insufficient SOE</td>
<td>Insufficient (Imprecise, inconsistent)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT²⁴⁴ (738)</td>
<td>Inconclusive: SOE was insufficient given evidence from 1 trial with moderate study limitations.</td>
<td>Insufficient (one study, moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs²³³,²⁴¹,²⁴⁴ (1,654)</td>
<td>No difference: No difference between oral agents with IUI versus gonadotropins with IUI</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT²³³ (742)</td>
<td>Increased risk: Greater multiple gestations with gonadotropins compared to either clomiphene or letrozole</td>
<td>Low (one study)</td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 RCTs²³⁴,²⁴¹ (765)</td>
<td>Inconclusive: SOE was insufficient given inconsistent and imprecise findings.</td>
<td>Insufficient (Inconsistent and Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Different Treatment Strategies for Controlled Ovarian Hyper</td>
<td>Live birth</td>
<td>3 RCTs²⁶⁹,²¹¹,²²³ (837)</td>
<td>Inconclusive: SOE was insufficient given inconsistent and imprecise findings.</td>
<td>Insufficient (Inconsistent and Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>---------------------------</td>
<td>------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>stimulation with Gonadotropins &amp; IUI</td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs¹¹¹,¹²¹,¹³⁰ (929)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence about the outcome from 3 studies targeting each an individual intervention.</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>1 RCT²²⁰ (161)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 trial with no events.</td>
<td>Insufficient (Imprecise, one small study)</td>
</tr>
<tr>
<td>Immediate IVF vs. Other Treatments Prior to IVF</td>
<td>Live birth</td>
<td>3 RCTs¹¹²,¹²¹,¹²⁰ (812)</td>
<td>No difference: Live birth does not differ between differing strategies of other treatments prior to IVF</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>2 RCTs¹¹²,¹²⁰ (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>3 RCTs¹¹²,¹²¹,¹²⁰ (812)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>3 RCTs¹¹²,¹²¹,¹²⁰ (812)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Death</td>
<td>2 RCTs¹¹²,¹²⁰ (657)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence for a rare event which requires a larger data set to draw inferences</td>
<td>Insufficient (Imprecise, rare events)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birthweight</td>
<td>2 RCTs¹¹²,¹²⁰ (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Time to pregnancy</td>
<td>2 RCTs¹¹²,¹²⁰ (657)</td>
<td>Reduction: Shorter time to pregnancy with immediate IVF compared with other treatments prior to IVF</td>
<td>Moderate</td>
</tr>
<tr>
<td></td>
<td>Costs</td>
<td>1 RCT³² (619)</td>
<td>Inconclusive: Insufficient SOE given one study with imprecise and overlapping findings</td>
<td>Insufficient (Imprecise, one study)</td>
</tr>
<tr>
<td></td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 RCTs¹¹²,¹²⁰ (657)</td>
<td>No difference: No significant difference between other treatments prior to IVF and immediate IVF.</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>Expectancy Management vs. Other Interventions</td>
<td>Live birth</td>
<td>1 RCT¹¹⁹ (161)</td>
<td>Inconclusive: Insufficient SOE given inconsistent findings from small studies</td>
<td>Insufficient (Inconsistent; small studies)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT²²⁵ (105)</td>
<td>Inconclusive: Insufficient SOE given imprecise findings from one small study</td>
<td>Insufficient (Imprecise findings from small study)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT²²⁵ (105)</td>
<td>Inconclusive: Insufficient SOE given imprecise findings from one small study</td>
<td>Insufficient (Imprecise findings from small study)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>---------------------------</td>
<td>------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Costs</td>
<td>1 RCT129 (253)</td>
<td>Inconclusive: Insufficient SOE given imprecise findings from one small study published outside US</td>
<td>Insufficient (Imprecise findings from small study)</td>
<td></td>
</tr>
<tr>
<td>ART: ICSI</td>
<td>Live birth</td>
<td>1 RCT214 (156)</td>
<td>Inconclusive: Insufficient SOE given imprecise findings from one study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT214 (156)</td>
<td>Inconclusive: Insufficient SOE given imprecise findings from one study with moderate risk of bias</td>
<td>Insufficient (Imprecise findings with moderate study limitations)</td>
</tr>
<tr>
<td>ART: IVF vs. ICSI</td>
<td>Costs</td>
<td>1 Obs215 (154)</td>
<td>Inconclusive: SOE was insufficient given imprecise evidence from 1 small observational trial.</td>
<td>Insufficient (Imprecise, one small study)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birth weight</td>
<td>1 Obs185 (90,401 cycles)</td>
<td>No difference: No significant differences in rates of low birth weight (OR 1.03, 95% CI 0.73 to 1.47) between ICSI versus conventional-IVF cycles</td>
<td>Low (one study with moderate study limitations)</td>
</tr>
<tr>
<td>ART: Unspecified</td>
<td>Long-term outcomes: Child (cancer)</td>
<td>1 Obs213 (33,840)</td>
<td>No difference: The overall cancer incidence was not elevated in children born after assisted conception for unexplained infertility. The SIR for all cancers in children conceived with assisted conception for unexplained infertility compared to the general population of the same age was 0.84 (95% CI, 0.57 to 1.18).</td>
<td>Low (Moderate study limitations)</td>
</tr>
</tbody>
</table>

*Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: AH=assisted hatching; AMA=advanced maternal age; ART=assisted reproductive technology; CC=clomiphene citrate; CI=confidence interval; DHEA=dehydroepiandrosterone; FASTT=Fast Track and Standard Treatment; FSH=follicle-stimulating hormone; hMG=human menopausal gonadotropin; HR=hazard ratio; ICER=incremental cost-effectiveness ratio; ICSI=intra-cytoplasmic sperm injection; IUI=intraruterine insemination; IVF=in vitro fertilization; KQ=Key Question; N=number of patients/participants; NA=not applicable; Obs=observational study; OHSS=ovarian hyperstimulation syndrome; OR=odds ratio; PGS=preimplantation genetic screening; RCT=randomized controlled trial; rFSH=recombinant follicle-stimulating hormone; RIF=recurrent implantation failure; SIR=standardized incidence ratio; uFSH=urinary follicle-stimulating hormone

**Key Question 4. Tubal and Peritoneal Factor Infertility**

KQ 4. What are the comparative safety and effectiveness of available treatment strategies for women with tubal or peritoneal factors (e.g., pelvic adhesions) who are infertile and who wish to become pregnant?

KQ 4a. Does the optimal treatment strategy vary by patient characteristics such as age, ovarian reserve, race, BMI, presence of other potential causes of female infertility, or presence of male factor infertility?
Description of Included Studies for KQ 4 (Tubal and Peritoneal Factor Infertility)

We identified eight individual studies\textsuperscript{123,184,185,248-252} that addressed outcomes after treatment for tubal or peritoneal factor infertility.

Among the eight studies that addressed outcomes after treatment for tubal or peritoneal factor infertility, three were RCTs. One was rated as good quality,\textsuperscript{250} one was rated as fair quality,\textsuperscript{252} and one was rated as poor quality.\textsuperscript{248} The remaining five studies were observational; three were good quality,\textsuperscript{184,249,251} and two were fair quality.\textsuperscript{123,185} Geographically, three studies were conducted in the U.K. or continental Europe\textsuperscript{184,249,250}, three were conducted in the U.S.\textsuperscript{123,185,251}, and two were conducted in Asia.\textsuperscript{248,252} Four studies were conducted in, or used data from, fertility subspecialty clinics,\textsuperscript{185,250-252} whereas setting was not specified for the remaining four.\textsuperscript{123,184,248,249} With regard to funding source, one study was government funded,\textsuperscript{252} one was industry funded,\textsuperscript{250} two reported non-government, non-industry funding,\textsuperscript{123,184} and, for the remaining four, funding source was not specified or was unclear.\textsuperscript{185,248,249,251}

Outcomes and complications of pregnancy were described in three RCTs\textsuperscript{248,250,252} and five observational studies.\textsuperscript{27,123,124,185,251} One observational study reported on long-term outcomes in the child.\textsuperscript{184} One observational study reported on treatment costs.\textsuperscript{249}

The main classes of treatment investigated were oral ovulation induction with or without IUI, gonadotropins with or without IUI, surgical treatment and surgery with hormonal adjunctive therapy, and ART (IVF or ICSI) alone or with adjunctive acupressure treatment. We did not perform meta-analysis because of the lack of studies reporting results for similar outcomes and treatment comparisons.

In addition to the above studies, one good-quality systematic review that addressed the comparative effectiveness of various treatments for infertility in women with tubal factor infertility (5 studies, 646 patients) is discussed below and the consistency of its findings with our included studies is incorporated in to our strength of evidence ratings.\textsuperscript{93}

Key Points for Tubal or Peritoneal Factor Infertility

Key findings for patients with tubal or peritoneal factor infertility included:

- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE)
- The live birth rate was lower in women undergoing ICSI as compared to conventional IVF (low SOE)
- There was no difference between type 1 diabetes mellitus diagnoses in children born to patients with tubal factor infertility conceived with ART compared to children conceived with no fertility treatment (moderate SOE)
- SOE was rated insufficient for all other comparisons/outcomes.

Detailed Synthesis by Treatment for Tubal or Peritoneal Factor Infertility

Included studies and their findings for the following treatments for tubal or peritoneal infertility are detailed in this section:

1. Surgical Management
2. ART
1. Surgical Management for Tubal or Peritoneal Factor Infertility

Three studies and one systematic review explored the use of surgical management strategies in women with tubal or peritoneal factor infertility.

We identified one RCT of a surgical intervention that examined live birth outcomes. A good-quality RCT compared hysteroscopic proximal tubal occlusion via Essure® to laparoscopic salpingectomy prior to IVF/ICSI in 85 women. They found that there was no difference in pregnancy complications of miscarriage or ectopic pregnancy between groups, but that women who underwent laparoscopic salpingectomy had a significantly higher live birth rate than women who underwent Essure® prior to IVF/ICSI in the ITT analysis (21.4% vs. 46.5%, respectively). However, this difference was no longer significant in the per protocol analyses (p = 0.143). Given imprecise findings from just one small study the SOE was rated as insufficient.

The risk of ectopic pregnancy was examined in one poor-quality RCT of women who underwent transcervical falloposcopy tubal dilatation as an adjunctive therapy during surgical treatment of tubal fertility. The study examined 468 infertile women with evidence of fallopian tube disease who underwent laparoscopic salpingolysis and/or salpingostomy. Patients were included if they had tubal factor infertility on hysterosalpingography, and were excluded if they had evidence of uterine fibroids, endometriosis, male factor infertility, genital malformations, unilateral salpingectomy, or endocrine pathology. A total of 256 patients were randomized to the intervention, transcervical falloposcopy tubal dilatation, and 212 patients did not receive the additional procedure (control group). The ectopic pregnancy rate in women who underwent transcervical falloposcopy tubal dilatation was 2% compared with 5.4% in the control group (p=0.647) (insufficient SOE).

One good-quality study evaluated costs of different diagnostic and treatment scenarios for women with tubal infertility. Six different scenarios involving different diagnostic approaches (no diagnostics, hysterosalpingogram, laparoscopy) and delayed or immediate IVF treatment were examined using a computer-generated Markov model. Costs were based on those in the Netherlands. Costs per live birth of the various diagnostic and treatment scenarios differed by more than 3000 euros. The costliest intervention was no diagnostics with immediate IVF treatment (8927 euros), and the least expensive was no diagnostics with 12 cycles of expectant management followed by 3 cycles of IVF (6459 euros) (insufficient SOE).

Finally, one good-quality Cochrane systematic review involving 5 RCTs with a total of 646 women explored surgical treatment for tubal disease in women undergoing IVF. None of the included trials reported on the primary outcome of live birth, and therefore they were excluded from our systematic review. They did, however, report an increase in ongoing pregnancy and clinical pregnancy with laparoscopic salpingectomy for hydrosalpinges prior to IVF versus non-surgical management. No significant differences were seen in any of the adverse effects of surgical treatments.

2. ART for Tubal or Peritoneal Factor Infertility

One RCT and four observational studies examined treatment outcomes in women with tubal factor infertility undergoing IVF/ICSI.

One fair quality RCT compared three different frequencies of transcutaneous electrical acupoint stimulation (TEAS) and a non-TEAS control group prior to IVF among 481 women with bilateral tubal blockage. Those assigned to one of the three TEAS arms received 30-minutes of TEAS 24 hours before transvaginal oocyte retrieval and 2 hours after embryo transfer. TEAS was administered at low frequency (2 Hz), high frequency (100 Hz), or
alternating low and high frequencies (2/100 Hz). The women receiving TEAS at alternating low and high frequencies 2/100Hz had a higher live birth rate (55/144, 48.25%) when compared to each of the other study arms (p<0.05).

Live birth rates was also evaluated in one observational study. The good-quality study251 used the Society for Assisted Reproductive Technology Clinic Outcome Reporting System (SART CORS) database to compare outcomes for conventional IVF and use of ICSI in women with tubal ligation. In adjusted analyses, use of ICSI was associated with lower odds of live birth (AOR 0.77, 0.69-0.85) (low SOE). There were no significant differences in plurality, mean length of gestation, or birth weight.251

An analysis of the SART CORS database compared success and complications associated with treatment patterns for women undergoing IVF/ICSI: live birth rates and maternal and neonatal complications based on the number of embryos transferred in 69,028 ART cycles.123 Among women with tubal infertility, the live birth rate after treatment with IVF or ICSI was 44.4 percent. In women with tubal fertility, the live birth rate was higher in couples who underwent two embryo transfer (47.5%) as compared with single-embryo transfer (41.1%) (p=0.05) (low SOE).123

Other outcomes of interest were evaluated in three observational studies. One fair-quality study185 used the SART CORS database to examine prevalence of preterm delivery and low birth weight among singletons conceived with ICSI compared to conventional IVF. In secondary analyses conducted on the subset of patients who used autologous sperm, had a favorable fertility prognosis based on female age <35, and > three oocytes retrieved, no significant differences in rates of preterm delivery or low birth weight was observed among those with tubal factor only infertility (preterm delivery OR 1.14, 95% CI -.79 to 1.65, p = .49; low birth weight OR 0.88, 95% CI 0.56 to 1.37, p = .56) (low SOE).185

One nationwide birth cohort study identified all pregnancies with a live-single born child over an 8-year period in Denmark and compared the incidence of type I diabetes among those conceived with fertility treatment to those conceived naturally.184 There was no association between tubal factor infertility as an indicator for fertility treatment and the subsequent development of Type I diabetes in offspring (HR 1.08, 95% to 0.61 to 1.91) (moderate SOE).

Strength of Evidence for Tubal and Peritoneal Factor Infertility

Table 25 summarizes the SOE for the findings described above. The SOE was judged to be insufficient for all outcomes primarily due to imprecision based on few studies meeting our inclusion criteria.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hysteroscopic proximal occlusion vs. laparoscopic salpingectomy</td>
<td>Live birth (patient)</td>
<td>1 RCTs250 (85)</td>
<td>Inconclusive. SOE was insufficient given imprecise findings from 1 small study</td>
<td>Insufficient (Imprecise, 1 small study)</td>
</tr>
<tr>
<td>Transcervical falloposcopy tubal dilatation vs. no intervention</td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT248 (468)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from 1 trial with high potential limitations</td>
<td>Insufficient (Imprecise findings with high study limitations, 1 study)</td>
</tr>
</tbody>
</table>
Comparison Outcome Study Design (Sample Size) Conclusion Strength of Evidence (Rationale)\(^a\)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hysterosalpingogram, laparoscopy, no intervention vs. IVF</td>
<td>Costs</td>
<td>1 Obs(^{129}) (NA)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from 1 modeling study</td>
<td>Insufficient (Imprecise; indirect findings, 1 study)</td>
</tr>
<tr>
<td>ART: 2-embryo transfer vs. 1-embryo transfer</td>
<td>Live birth (patient)</td>
<td>1 Obs(^{133}) (69,028 cycles)</td>
<td>Improvement. The live birth rate per cycle was higher in couples who underwent 2 embryo transfer as compared with single-embryo transfer</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>ART: Frequencies of transcutaneous electrical acupoint stimulation prior to IVF</td>
<td>Live birth (patient)</td>
<td>1 RCT(^{252}) (481)</td>
<td>Inconclusive. SOE was insufficient given imprecise findings from 1 fair-quality study</td>
<td>Insufficient (Imprecise, 1 fair-quality study)</td>
</tr>
<tr>
<td>ART: IVF vs. ICSI</td>
<td>Live birth (patient)</td>
<td>1 Obs(^{251}) (7145)</td>
<td>Reduction: Lower odds of live birth with ICSI</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td>ART: IVF+ICSI vs. IVF</td>
<td>Neonatal outcomes: Birth weight</td>
<td>1 Obs(^{183}) (90,401 cycles)</td>
<td>No difference: No significant differences in rates of low birth weight (OR 0.88, 95% CI 0.56 to 1.37) between ICSI versus conventional-IVF cycles</td>
<td>Low (1 study with moderate study limitations)</td>
</tr>
<tr>
<td>ART vs. no fertility treatment</td>
<td>Long-term outcomes: Child (type 1 diabetes mellitus)</td>
<td>1 Obs(^{184}) (565,116 pregnancies)</td>
<td>No difference: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with tubal factor infertility conceived with ART compared to children conceived with no fertility treatment</td>
<td>Moderate (Imprecise)</td>
</tr>
</tbody>
</table>

\(^a\)Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: ART=assisted reproductive technology; CI=confidence interval; HR=hazard ratio; ICSI=intracytoplasmic sperm injection; IUI=intruterine insemination; IVF=in vitro fertilization; KQ=Key Question; N=number of patients/participants; NA=not applicable; Obs=observational study; RCT=randomized controlled trial; RR=relative risk; SIR=standardized incidence ratio

**Key Question 5. Male Factor Infertility**

KQ 5. What are the comparative safety and effectiveness of available treatment strategies for couples with male factor infertility and no evidence of an underlying diagnosis associated with infertility in the female partner?

KQ 5a. Does the optimal treatment strategy vary by characteristics in either partner such as age, ovarian reserve, race, or BMI?
Description of Included Studies for KQ 5 (Male Factor)

We identified 23 individual studies\textsuperscript{119,123-125,184-186,208,213,221,236,242,253-264} that addressed the comparative effectiveness or safety of interventions applied to patients with male factor infertility.

Eleven studies were RCTs\textsuperscript{119,208,221,242,253,255,256,258,261,262,264} and 12 were observational studies.\textsuperscript{123,124,184-186,213,236,254,257,259,260,263} Of the 11 RCTs, four studies were rated good quality,\textsuperscript{119,208,261,262} five were rated fair quality,\textsuperscript{221,242,256,258,264} and two were rated poor quality.\textsuperscript{253,255} Of the 12 observational studies, three were rated good quality,\textsuperscript{184,186,236} seven were rated fair quality,\textsuperscript{123,124,185,213,254,260,263} and two were rated poor quality.\textsuperscript{257,259}

Of the 23 studies, six were conducted in the United States,\textsuperscript{123,124,185,236,254,263} 10 in the UK or continental Europe,\textsuperscript{119,184,186,213,221,242,253,256,259,264} five in the Middle East,\textsuperscript{208,255,258,260,261} and two in Asia.\textsuperscript{257,262} All but three studies were conducted in subspecialty practices; the remaining three did not report the setting or the setting was unclear.\textsuperscript{123,184,213} Finally, six studies reported government funding,\textsuperscript{119,213,254,256,259,265} one reported industry funding,\textsuperscript{186} three studies reported non-government, non-industry funding,\textsuperscript{123,184,257} two studies reported a combination of funding sources,\textsuperscript{124,236} and the remaining eleven studies did not report funding source or the funding source was unclear.\textsuperscript{185,208,221,242,253,255,258,260,261,263,264} The interventions and comparisons evaluated in the included studies are summarized in Appendix E.

In addition to the above studies, four good-quality systematic reviews\textsuperscript{72,79,265,266} that addressed the comparative effectiveness of various treatments in men with male factor infertility are discussed below and the consistency of their findings with our included studies are incorporated in to our strength of evidence ratings.

Key Points for Male Factor Infertility

Key findings for patients with male factor infertility included:

- Live birth rate (moderate SOE) and miscarriage (low SOE) did not differ between intracytoplasmic sperm injection (ICSI) and intracytoplasmic morphological sperm injection (IMSI). Of note, IMSI is not used in the United States.
- There was no difference in live birth rates or any adverse pregnancy events between couples using frozen embryo versus fresh embryo transfer (low SOE).
- The overall cancer incidence was not elevated in children born after assisted conception for male factor infertility (low SOE).
- There was no difference between type 1 diabetes mellitus diagnoses in children born to patients with male factor infertility conceived with ART compared to children conceived with no fertility treatment (moderate SOE).
- Live birth rate (low SOE) improved with vitamin E or zinc supplementation relative to placebo or no supplementation.
- As with other indications for IVF, use of single-embryo transfer is associated with slightly lower live birth rates but significantly reduced multiple gestation rates (low SOE).
Detailed Synthesis by Treatment for Male Factor Infertility

Included studies and their findings for the following treatments for unexplained infertility are detailed in this section:

1. IUI With Adjunct Treatment for Female Partners
2. ART
   a. IVF
   b. ICSI
   c. IVF Versus ICSI
   d. ART Unspecified
3. Other Strategies

1. IUI With Adjunct Treatments for Female Partners With Male Factor Infertility

A single fair-quality RCT\textsuperscript{221} that compared single versus double IUI in multifollicular ovarian hyperstimulation cycles reported live birth rates and pregnancy complications. Clomiphene citrate or recombinant FSH or both were used for ovarian hyperstimulation in both study arms. Of the 228 women in the trial, male factor was the cause of infertility for 67 couples in the single IUI arm and 65 couples in the double IUI arm. Among the couples with male factor infertility, double IUI was associated with a crude OR of 0.71 (95% CI, 0.21 to 2.37; p=0.764) relative to single IUI for live birth rate. This trial was conducted in a single subspecialty clinic in Turkey and given the imprecise findings was rated as insufficient strength of evidence.

A single good-quality RCT\textsuperscript{208} compared rFSH with clomiphene citrate in IUI cycles in couples with infertility. Of the 219 couples in the trial, male factor was the cause of infertility for 15 couples in rFSH group and 22 couples in the clomiphene citrate group. Among the couples with male factor infertility, the live birth rate per patient was 20 percent in the rFSH group and 13.6 percent in the clomiphene citrate group (p value reported as “not significant”). This trial was conducted in a single subspecialty clinic in Turkey (insufficient SOE).

Related to these included studies, two good-quality systematic reviews published in 2007\textsuperscript{265} and 2017\textsuperscript{266} evaluated outcomes associated with IUI for male-factor infertility. The study by Bensdorp, 2007 compared ovarian hyperstimulation with IUI with IUI alone among couples with male factor infertility did not demonstrate a significant difference in live birth rates (OR 0.87; 95% CI, 0.28 to 2.70). The more recent systematic review by Cissen, 2016 also compared ovarian hyperstimulation with IUI with IUI alone among couples with male factor infertility summarizing 3 RCTs of a total of 346 couples (OR 1.34, 95% CI 0.77, 2.33; I\textsuperscript{2}=0%) findings no evidence of a difference with live births. They noted very low quality of evidence for this outcome given potential high risk of bias in the included studies, inconsistencies in the data, and imprecise findings. This review also identified two RCTs (1 low quality and 1 very low quality) examining IVF versus IUI in natural cycles or cycles with ovarian hyperstimulation. This review found no evidence of a difference in live birth rates between IVF versus IUI in natural cycles (1 RCT, 53 couples; OR 0.77, 95% CI 0.25, 2.53) or IVF versus IUI in cycles with ovarian hyperstimulation (2 RCTs, 86 couples; OR 1.03 95% CI 0.43, 2.45; I\textsuperscript{2}=0%) and again with very low quality evidence.
2. ART for Male Factor Infertility

IVF for Male Factor Infertility

Three studies reported findings from comparisons of varying IVF procedures for patients with male factor infertility.\textsuperscript{123,257,262} Findings for these studies are summarized in Table 26. We identified one good-quality RCT that investigated fresh versus frozen embryo transfer on multiple outcomes including live birth rates, multiples, birth weight, ectopic pregnancy, and congenital anomalies.\textsuperscript{262} Patients within this RCT had either tubal or male factor infertility and the results were not stratified by type. Approximately 40 percent of the population had male factor or both male and tubal factor infertility. The strength of evidence for the findings from this trial was reduced given the inclusion of couples with tubal factor infertility in the cohort.

Two additional studies reported live birth rate per cycle,\textsuperscript{123,257}, one of which also reports neonatal outcomes and congenital anomalies.\textsuperscript{257}

Table 26. Outcomes for comparisons of IVF in couples with male infertility

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Study Design (N Patients)</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Results Intervention N (%) (95% CI)</th>
<th>Results Comparator N (%) (95% CI)</th>
<th>P value</th>
<th>Summary of Study Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Live birth: Any/patient</td>
<td>Tsai, 2011\textsuperscript{257} RCT (191 cycles)</td>
<td>TESE</td>
<td>Ejaculated OAT</td>
<td>40.7% (95% CI)</td>
<td>30.2% (95% CI)</td>
<td>0.197</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Luke, 2010\textsuperscript{123} RCT (69,028 ART cycles)</td>
<td>ICSI or assisted hatching (1 embryo transferred)</td>
<td>ICSI or assisted hatching (multiple embryos transferred)</td>
<td>44.7% (95% CI)</td>
<td>52.1% (2 embryos) 46.9% (3 embryos) 40.4% (4 embryos) (number transferred)</td>
<td>&lt;0.001</td>
<td>Greater live births with multiple embryos transferred compared to 1 embryo transferred</td>
</tr>
<tr>
<td></td>
<td>Shi, 2018\textsuperscript{262} RCT (2,157 patients)</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>48.7% (95% CI)</td>
<td>50.2% (95% CI)</td>
<td>0.50</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>Shi, 2018\textsuperscript{262} RCT (2,157 patients)</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>2.7% (95% CI)</td>
<td>1.7% (95% CI)</td>
<td>0.23</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>Shi, 2018\textsuperscript{262} RCT (2,157 patients)</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>Among clinical pregnancies 9.4%</td>
<td>Among clinical pregnancies 11.5%</td>
<td>0.22</td>
<td>No difference</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>Shi, 2018\textsuperscript{262} RCT (2,157 patients)</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>Singleton 31.3% Twin 17.4% Triplet 0.1%</td>
<td>Singleton 34.0% Twin 16.0% Triplet 0.2%</td>
<td>0.18</td>
<td>0.40 1.00 No difference</td>
</tr>
<tr>
<td>Outcome</td>
<td>Study Design (N Patients)</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Results Intervention N (%) (95% CI)</td>
<td>Results Comparator N (%) (95% CI)</td>
<td>P value</td>
<td>Summary of Study Findings</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>--------------------------</td>
<td>--------------</td>
<td>----------------------------</td>
<td>------------------------------------</td>
<td>------------------------------------</td>
<td>---------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>Tsai, 2011^25^3</td>
<td>TESE</td>
<td>Ejaculated extreme severe OAT sperm</td>
<td>38.3%</td>
<td>23.8%</td>
<td>0.292</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Shi, 2018^26^2</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>3373g (+/-515)</td>
<td>3380g (+/-502)</td>
<td>0.85</td>
<td>No difference</td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>Tsai, 2011^25^3</td>
<td>TESE</td>
<td>Ejaculated extreme severe OAT sperm</td>
<td>1.7% (major) 8.3% (minor)</td>
<td>4.7% (major) 4.8% (minor)</td>
<td>0.454</td>
<td>No difference</td>
</tr>
<tr>
<td></td>
<td>Shi, 2018^26^2</td>
<td>Frozen embryo transfer</td>
<td>Fresh embryo transfer</td>
<td>2.2%</td>
<td>3.6%</td>
<td>0.12</td>
<td>No difference</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; HR=hazard ratio; ICSI=intracytoplasmic sperm injection; N=number of patients; NS=not statistically significant; OAT=oligo-astheno-teratozoospermia; RCT=randomized control trial; SOE=strength of evidence; TESE=extracted testicular sperm

ICSI for Male Factor Infertility

Two RCTs were identified for analyzing outcomes of male factor infertility while undergoing variations of ICSI^25^5,26^4. One fair-quality RCT investigated the effects of rFSH and rLH versus rFSH alone on miscarriage in couples with repeated implantation failures. The study found a significantly lower miscarriage rate in the rFSH and rLH group (21%) as compared to the rFSH alone group (37.5%; p<.01).^26^4 Strength of evidence was rated as insufficient given findings from one small study with potential limitations.

The second RCT was rated as poor-quality.^25^5 This study was conducted in Iran and had randomly allocated 182 couples undergoing ICSI for male factor infertility to laser assisted hatching (n=90) or intact embryo transfer without laser assisted hatching (n=92). The live birth rate per patient was 11.1% in the laser assisted hatching arm and 8.6% in the control arm (p=0.06). This study reported 2 sets of twins among the 10 live births in the laser assisted hatching arm and 2 sets of twins among the 8 live births in the conventional ICSI arm (OR 0.6; 95% CI, 0.08 to 6.9). In addition, one congenital anomaly was reported among the 10 live births in the laser assisted hatching arm and zero congenital anomalies among the 8 live births in the conventional ICSI arm (OR 0.9; 95% CI, 0.07 to 1.1) (insufficient SOE for all outcomes).^25^5

IVF Versus ICSI for Male Factor Infertility

Live Birth

Three RCTs (2 fair quality^25^6,25^8 and 1 poor quality^25^3) and 2 fair-quality observational studies^12^4,25^4 reported live birth rates for comparisons between IVF and ICSI for couples with male factor infertility.

We conducted a meta-analysis of the live birth rates per cycle reported by the 3 RCTs summarized above that compared ICSI with IMSI (Figure 3). The summary estimate of the OR for live birth per cycle associated with ICSI was 1.218 (95% CI, 0.779 to 1.903) using a fixed
effect model (with identical results using a random effects model). The Q-value was 0.2913 with 2 degrees of freedom, with an $I^2$ of 0%, suggesting no significant heterogeneity. This meta-analysis corroborates the findings of each of the three primary studies, which is that there is no significant difference in live birth rates per cycle between ICSI and IMSI procedures for male factor infertility. The systematic review published in 2013\textsuperscript{79} included 9 trials with 2,014 couples however only the study by Balaban and colleagues\textsuperscript{258} reported on live birth. We rated the strength of evidence for no difference in live birth as moderate based on the findings of our meta-analysis.

**Figure 3. Effect of ICSI versus IMSI on live birth**

<table>
<thead>
<tr>
<th>Study name</th>
<th>Odds ratio</th>
<th>Lower limit</th>
<th>Upper limit</th>
<th>Odds ratio and 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Sala, 2015</td>
<td>1.301</td>
<td>0.494</td>
<td>3.423</td>
<td></td>
</tr>
<tr>
<td>Leandri, 2013</td>
<td>1.102</td>
<td>0.611</td>
<td>1.990</td>
<td></td>
</tr>
<tr>
<td>Balaban, 2011</td>
<td>1.485</td>
<td>0.569</td>
<td>3.876</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>1.218</td>
<td>0.779</td>
<td>1.903</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; ICSI=intracytoplasmic sperm injection; IMSI=intracytoplasmic morphologically selected sperm injection

We also identified two fair-quality observational studies\textsuperscript{124,254} that analyzed data from the SART CORS database to compare ICSI with IVF for male factor infertility. These two studies may report data from the same patients. One of the studies compared live birth rates of both couples with male factor infertility who received conventional IVF (N=4973 cycles) or combined IVF with ICSI (N=72,459 cycles).\textsuperscript{124} Among couples with male factor infertility, the live birth rate for the 1989 cycles that resulted in a pregnancy after treatment with conventional IVF was 84.9%, compared with 85.2% live birth rate among the approximately 32,896 cycles that resulted in a pregnancy after treatment with ICSI. This corresponds to an OR of 0.98 (95% CI, 0.90 to 1.06) for live birth rate given a pregnancy associated with IVF relative to ICSI among couples with male factor infertility and supports the findings from our meta-analysis. The second study compared outcomes associated with 10,933 cycles of conventional IVF with 153,968 ICSI cycles for male factor infertility reported to the SART registry during 1996-2012.\textsuperscript{254} The adjusted RR for live birth per cycle among transfers associated with conventional IVF relative to ICSI was 0.98 (95% CI, 0.95 to 1.02; p>0.99). The multivariate analysis was adjusted for maternal age, number of prior live births, number of prior miscarriages, number of prior assisted reproductive technology cycles, number of oocytes retrieved, number of embryos cryopreserved, donor egg/embryo, donor sperm, use of preimplantation genetic testing, and infertility diagnosis (tubal factor, endometriosis, uterine factor, ovulatory disorder, and diminished ovarian reserve).
this observational study supports the findings from our meta analysis of no difference between IVF and ICSI for the outcome of live birth (moderate SOE).

**Miscarriage**

We identified two studies (one poor-quality RCT\(^\text{253}\) and one fair-quality observational study\(^\text{254}\)) that compared ICSI to IMSI for male factor infertility in relation to miscarriage rates. The RCT\(^\text{253}\) reported miscarriage rates of 20.5% (15/73) for ICSI and 22.9% (11/48) for IMSI (p=0.823) while the observational study\(^\text{254}\) calculated an adjusted RR of 0.97 (95% CI, 0.91 to 1.04) for miscarriage among pregnancies for ICSI relative to IVF among couples with male factor infertility.

Our findings are consistent with the good-quality systematic review by Teixeira and colleagues which consisted of 6 RCTs (552 women)\(^\text{79}\) that concluded that there is no evidence of effect on miscarriage between ICSI and IMSI. Although all of these studies support there being no difference in miscarriage rates, the SOE was rated as low given large imprecision in the findings and the high risk of bias in the included studies.

**Multiple Births**

In addition to miscarriage rates, the two studies above\(^\text{253,254}\) reviewed multiple birth events when comparing ICSI to IMSI for male factor infertility. The poor-quality RCT reported that 22.2% (2/9) and 36.4 % (4/11) of births were twin deliveries in the ICSI and IMSI arms, respectively (p=0.496).\(^\text{253}\) The fair-quality observational study by Boulet and associates\(^\text{254}\) compared outcomes associated with conventional IVF with ICSI calculated an adjusted RR of 0.87 (95% CI, 0.83 to 0.91) for multiple live births among pregnancies for ICSI relative to IVF among couples with male factor infertility. Given the quality of these two studies and the imprecision of the findings the SOE was rated as insufficient.

**Birthweight**

One poor-quality RCT\(^\text{253}\) and three fair-quality observational studies compared birthweight for ICSI versus conventional IVF.\(^\text{124,185,254}\) Boulet’s observational study\(^\text{254}\) compared birthweight for ICSI (4230 live births) versus conventional IVF (60,273 live births). Among couples with male factor infertility, they calculated an adjusted RR of low birthweight of 0.93 (95% CI, 0.88 to 0.98) for ICSI relative to IVF. The second observational study\(^\text{124}\) evaluated outcomes associated with conventional IVF versus ICSI among both couples with male factor infertility reported singleton birth weights of 3270 grams (standard deviation [SD] 580) among conventional IVF recipients versus 3266 grams (SD 637) among ICSI recipients (p=0.67). The third fair-quality observational study\(^\text{185}\) examined neonatal outcomes for fresh ICSI compared with IVF cycles among patients with male factor infertility. This study used data from the society for assisted reproductive technologies clinical outcomes reporting system database from 2004 to 2013. Conventional IVF participants were one-to-one propensity score matched with ICSI participants. Among matched patients with male factor infertility (n=2,184) no significant association was found with preterm delivery (OR 0.77 95% CI 0.54, 1.12) or low birth weight (OR 1.06 95% CI 0.70, 1.61).

Finally, one poor-quality RCT found that the mean birthweight was 2535 grams (SD 710) associated with IMSI versus 2789 grams (SD 575) associated with ICSI (p=0.492), not demonstrating a difference.\(^\text{253}\) Given the poor- and fair-quality of the included studies and the lack of quality RCT evidence, the SOE was rated as low for no difference in birthweight.
Congenital Anomalies

Three studies explore congenital anomalies associated with infertility treatments for male infertility. One poor-quality RCT looked at congenital anomalies when comparing ICSI to IMSI for male factor infertility. The study reported zero cases of congenital anomalies for ICSI and 18.2% (2/11) for IMSI (p=0.190). Additionally, a fair-quality observational study investigated rates of birth defects in patients who conceived through IVF and ICSI. This study found birth defects in 2.4% of infants conceived through IVF for male factor infertility, which was significantly lower than the 3.2% of infants with birth defects conceived through ICSI. However, these results should be interpreted with caution. The male infertility factor group in this study seems to include couples with and without female infertility indication.

Finally, a third fair-quality observational study was conducted in Israel in a single outpatient infertility clinic of couples presenting with male factor infertility. The objective of the study was to determine the birth defect rates among ICSI and IMSI groups. Among 1,981 pregnancies, sixty-three pregnancies involving fetal malformation were reported with 3.9% in the ICSI group and 2.3% in the IMSI group. However, no significant association was found between the two groups (OR 0.71, 95% CI 0.39–1.22). Given the study limitations of these included studies with imprecise findings, the SOE was rated as insufficient.

ART Unspecified for Male Factor Infertility

Live Birth Rates

We identified one good-quality RCT and one good-quality observational study that investigated live birth rates in couples undergoing IVF with male factor infertility. The RCT investigated the effects of a female weight-loss intervention on live birth rates in patients with both female and male factor infertility. The weight-loss intervention resulted in significantly more natural conceptions (rate ratio of intervention = 1.61; 95% CI 1.16 to 2.24), but significantly less live births of a healthy singleton at term than the control group (rate ratio of intervention = 0.77; 95% CI 0.60 to 0.99). These results were not stratified by infertility diagnosis (i.e. included both male and female factor infertility), however results reporting live birth rates by method of conception suggest that the female weight loss intervention had no effect on live birth rates for patients with male factor infertility. This is because the proportion of women who underwent ovulation induction (suggesting female factor infertility) was 26% versus 40% in intervention versus control (95% CI of RR does not cross 1.0), 24% versus 26% for IUI (primarily male factor; 95% CI of RR crosses 1.0) and 20% versus 28% for IVF/ICSI (combined female and male factor infertility; 95% CI just crosses 1.0).

The observational national cohort study from the Danish ART Registry looked at live birth rates in couples initiating fertility treatment with ART or IUI, and continuing to undergo either fertility treatment (regardless of initial treatment) for two years. Among male factor infertility patients (females<35 years old) whose first treatment was ART, two-year total live birthrates were 60.6% (across interventions: 56.3% for ART conception, 0.3% for IUI, and 5.0% for natural conception). Among male factor infertility patients (females≥35 years old) whose first treatment was ART, two-year birthrates were 37.3% for ART conception, 0.4% for IUI, and 5.2% for natural conception. For male factor infertility patients whose first treatment was IUI (females<35 years old), two-year birthrates were 20.2% for ART conception, 31.9% for IUI, and 8.5% for natural conception. For male factor infertility patients whose first treatment was IUI (females≥35 years old), two-year birthrates were 12.9% for ART conception, 25.1% for IUI, and 7.4% for natural conception. These results provide age-stratified probabilities of live birth for
patients with male factor infertility initiating with either ART or IUI and continuing to undergo either fertility treatment for 2 years.\textsuperscript{186}

**Birthweight**

One good-quality observational study\textsuperscript{236} looked at all singleton live births in Florida and Massachusetts between 2000 and 2010 and Michigan between 2000-2009 and used data from the States Monitoring Assisted Reproductive Technology (SMART) Collaborative. Both conventional and discordant-sibling pairs analyses were conducted. For the discordant sibling pairs the study population was restricted to singleton live births where one sibling was conceived through ART and the other was conceived naturally. In conventional analysis, when compared to naturally conceiving, those with male factor only indication for ART had significant associations with low birth weight (OR 1.15 95% CI 1.04 to 1.27) and preterm birth (OR 1.19 95% CI 1.10 to 1.29).\textsuperscript{236}

**Long-Term Outcomes: Child**

Risk for cancer in children conceived after IVF or ICSI was evaluated in a fair-quality cohort study of 106,013 children born in Britain between 1992 and 2008.\textsuperscript{213} For 24,427 of these children, the cause of infertility was male factor only. The average duration of follow-up was 6.6 years, with a maximum of 15 years of follow-up. Cancer diagnoses were identified through the National Registry of Childhood Tumors. Cancer was diagnosed in 16 children born after assisted conception for male factor infertility. The standardized incidence ratio (SIR) for all cancers in children conceived with assisted conception where the cause of infertility was male factor only compared to the general population of the same age was 0.92 (95% CI 0.53 to 1.49)

Despite the large number of children in the cohort, the overall number of cancers was small. Conclusions about the risk are tenuous due to the inability to examine risk for individual cancer types by underlying cause of infertility (low SOE).

In addition to cancer, we identified one good-quality Danish cohort study that evaluated the risk of type 1 diabetes mellitus in children conceived from ART for diagnosed male infertility. There was no significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with male factor infertility conceived with ART compared to children conceived with no fertility treatment (adjusted HR=0.83, 95% CI 0.55 to 1.46) (moderate SOE).\textsuperscript{184}

### 3. Other Strategies

**Exercise for Male Factor Infertility**

We identified one good-quality RCT that studied the effect of an exercise intervention for men with male factor infertility. Out of 197 patients in the exercise arm, there were 139 pregnancies and 127 live births, and in the non-exercise arm of 189 patients there were 5 pregnancies and 0 live births. Using the non-exercise arm as the reference value the authors calculated an odds ratio of live birth of 197.0 (95% CI 5.9-2149.6).\textsuperscript{261} Given imprecise findings from one small study the SOE was rated as insufficient.

**Antioxidants for Male Factor Infertility**

We did not identify any eligible studies that evaluated the comparative effectiveness of antioxidants on live births for couples with male factor infertility. We did, however, identify a good-quality systematic review\textsuperscript{72} that conducted a meta-analysis with data from 4 small RCTs, 3 of which were published prior to 2007; this meta-analysis demonstrated a significant increase in
live birth rate associated with vitamin E or zinc supplementation relative to placebo or no supplementation (OR 4.21; 95% CI, 2.08 to 8.51). However, these results are based only 44 live births out of 277 couples, and the overall quality of evidence within the included studies was rated low.

**Strength of Evidence for Male Factor Infertility**

Table 27 summarizes the SOE for the findings described above. The SOE was judged to be insufficient or low for all outcomes except for the comparison of IVF versus ICSI for live birth.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design and Sample Size</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single vs. double IUI</td>
<td>Live birth</td>
<td>2 RCTs&lt;sup&gt;208,221&lt;/sup&gt; (447 patients)</td>
<td><strong>Inconclusive.</strong> SOE was insufficient given imprecise evidence studies with varying quality and differing strategies.</td>
<td>Insufficient (Imprecise; findings from studies with varying quality and differing strategies)</td>
</tr>
<tr>
<td>rFSH vs. clomiphene citrate</td>
<td></td>
<td>2 SRs&lt;sup&gt;265,266&lt;/sup&gt; (4 studies, 1,278 couples)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IUI vs. IUI with ovarian hyperstimulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TESE vs. ejaculated OAT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ART IVF: ICSI or assisted hatching (1 embryo transferred) vs. ICSI or assisted hatching (multiple embryos transferred)</td>
<td>Live birth</td>
<td>2 Obs&lt;sup&gt;253,257&lt;/sup&gt; (272,717 cycles)</td>
<td><strong>Improv.</strong> Greater live births with multiple embryos transferred compared to 1 embryo transferred</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td>ART IVF: TESE vs. ejaculated OAT</td>
<td>Neonatal outcomes: Birthweight</td>
<td>1 Obs&lt;sup&gt;257&lt;/sup&gt; (191 cycles)</td>
<td><strong>Inconclusive.</strong> SOE was insufficient given imprecise evidence from one observational study with potential limitations</td>
<td>Insufficient (Imprecise findings with high study limitations)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>1 Obs&lt;sup&gt;257&lt;/sup&gt; (191 cycles)</td>
<td><strong>Inconclusive.</strong> SOE was insufficient given imprecise evidence from one observational study with potential limitations</td>
<td>Insufficient (Imprecise findings with high study limitations)</td>
</tr>
<tr>
<td>ART IVF: Frozen vs. fresh embryo transfer</td>
<td>Live birth</td>
<td>1 RCT&lt;sup&gt;262&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference:</strong> no difference in live birth rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Ectopic pregnancy</td>
<td>1 RCT&lt;sup&gt;262&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference:</strong> no difference in ectopic pregnancy rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design and Sample Size</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>----------------------------------</td>
<td>------------------------------</td>
<td>-----------------------------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;252&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in multiple birth rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;252&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in miscarriage rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT&lt;sup&gt;252&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in low birthweight rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>1 RCT&lt;sup&gt;252&lt;/sup&gt; (2,157 patients)</td>
<td><strong>No difference</strong>: no difference in congenital anomalies rates between couples using frozen embryo versus fresh embryo transfer</td>
<td>Low (one study, heterogeneous infertility indication)</td>
<td></td>
</tr>
<tr>
<td>ART ICSI: ICSI vs. Laser-assisted hatching</td>
<td>Live birth</td>
<td>1 RCT&lt;sup&gt;255&lt;/sup&gt; (182 patients)</td>
<td><strong>Inconclusive. SOE was insufficient given imprecise evidence from 1 low-quality trial</strong></td>
<td>Insufficient (Imprecise findings with high study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;254&lt;/sup&gt; (62)</td>
<td><strong>Inconclusive. SOE was insufficient given imprecise evidence from 1 small moderate-quality trial</strong></td>
<td>Insufficient (Imprecise findings from one small study with moderate limitations)</td>
<td></td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT&lt;sup&gt;255&lt;/sup&gt; (182 patients)</td>
<td><strong>Inconclusive. SOE was insufficient given imprecise evidence from 1 low-quality trial.</strong></td>
<td>Insufficient (Imprecise findings with high study limitations)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>1 RCT&lt;sup&gt;255&lt;/sup&gt; (182 patients)</td>
<td><strong>Inconclusive. SOE was insufficient given imprecise evidence from 1 low-quality trial.</strong></td>
<td>Insufficient (Imprecise findings with high study limitations and a suspected reporting bias)</td>
<td></td>
</tr>
<tr>
<td>IVF vs. ICSI</td>
<td>Live birth</td>
<td>3 RCTs&lt;sup&gt;253,256,258&lt;/sup&gt; (497 patients)</td>
<td><strong>No difference. Meta-analysis of 3 RCTs produced a summary estimate of the OR for live birth per cycle associated with ICSI of 1.218 (95% CI, 0.779 to 1.903) relative to IMSI and therefore does not demonstrate a difference between ICSI and IMSI.</strong></td>
<td>Moderate (Moderate study limitations)</td>
</tr>
<tr>
<td>Pregnancy complications: Miscarriage</td>
<td>1 RCT&lt;sup&gt;253&lt;/sup&gt; (121 patients)</td>
<td><strong>No difference. Both included studies and an existing systematic review supported no difference in miscarriage. SOE was reduced because of quality of included studies and imprecision of findings.</strong></td>
<td>Low (High study limitations, imprecise)</td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design and Sample Size</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>------------------------------</td>
<td>------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Pregnancy complications: Multiple births</td>
<td>1 RCT(^{233}) (121 patients) 1 Obs(^{24,4} ) (499,135 cycles)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from 1 trial and 1 observational study with moderate study limitations</td>
<td>Insufficient (Imprecise with moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Birthweight</td>
<td>1 RCT(^{233}) (121 patients) 3 Obs(^{24,185,254}) (862,062 cycles)</td>
<td><strong>No difference</strong>: No significant differences in rates of low birth weight between ICSI versus conventional-IVF cycles</td>
<td>Low (moderate study limitations)</td>
<td></td>
</tr>
<tr>
<td>Neonatal outcomes: Congenital anomalies</td>
<td>1 RCT(^{233}) (121 patients) 2 Obs(^{260,263}) (143,436)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from studies with significant study limitations.</td>
<td>Insufficient (Imprecise; findings with high study limitations)</td>
<td></td>
</tr>
<tr>
<td>ART unspecified</td>
<td>Long-term outcomes: Child (cancer)</td>
<td>1 Obs(^{113}) (924,427 patients)</td>
<td><strong>No difference</strong>: The overall cancer incidence was not elevated in children born after assisted conception for male factor infertility. The SIR for all cancers in children conceived with assisted conception for male factor infertility compared to the general population of the same age was 0.92 (95% CI, 0.53 to 1.49).</td>
<td>Low (Moderate study limitations)</td>
</tr>
<tr>
<td>Long-term outcomes: Child (type 1 diabetes mellitus)</td>
<td>1 Obs(^{184}) (565,116 pregnancies)</td>
<td><strong>No difference</strong>: No significant difference found between type 1 diabetes mellitus diagnoses in children born to patients with male factor infertility conceived with ART compared to children conceived with no fertility treatment</td>
<td>Moderate (Imprecise)</td>
<td></td>
</tr>
<tr>
<td>Other strategies: Exercise for Male Infertility</td>
<td>Live birth</td>
<td>1 RCT(^{261}) (386 patients)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from 1 small trial.</td>
<td>Insufficient (Imprecise; one study)</td>
</tr>
<tr>
<td>Other strategies: Antioxidant use for Male Infertility</td>
<td>Live birth</td>
<td>1 SR(^{272}) (4 studies of 277 couples)</td>
<td><strong>Improvement</strong>: Increase in live birth rate associated with vitamin E or zinc supplementation relative to placebo or no supplementation</td>
<td>Low (Imprecise, small studies)</td>
</tr>
</tbody>
</table>

*Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: ART=assisted reproductive technology; CI=confidence interval; HR=hazard ratio; ICSI=intracytoplasmic sperm injection; IMSI=intracytoplasmic morphologically selected sperm injection; IUI=intruterine insemination; IVF=**in vitro** fertilization; KQ=Key Question; N=number of patients/participants; NA=not applicable; NR=not reported; OAT=oligo-astheno-teratozoospermia; Obs=observational study; OR=odds ratio; RCT=randomized controlled trial; rFSH=recombinant follicle-stimulating hormone; RR=relative risk; SD=standard deviation; SIR=standardized incidence ratio; TESE=extracted testicular sperm
Key Question 6. Donors in Infertility

KQ 6. What are the short- and long-term health outcomes of donors in infertility?

KQ 6a. For female oocyte donors:
1. Do specific aspects of the pre-donation evaluation identify potential donors at greater risk for short- or long-term adverse outcomes (e.g., OHSS, quality-of-life issues)?
2. Do short- and long-term outcomes differ among different stimulation/retrieval protocols?

KQ 6b. For male semen donors:
1. Are there long-term health, quality-of-life, or other adverse outcomes associated with donation?

Description of Included Studies for KQ 6 (Donors)

We identified one fair-quality RCT267 and four retrospective observational studies, 3 fair-268-270 and 1 poor-quality,271 that addressed short- or long-term health outcomes of donors in infertility. Two of the studies were conducted in the United States.270,271 The remaining three studies were conducted in Europe.267-269 Four studies were conducted in a subspecialty practice, and the remaining study did not report the setting or the setting was unclear.271 None of the studies reported a funding source. We did not perform meta-analysis because of the lack of studies reporting results for similar outcomes and treatment comparisons. All of the included studies focused on oocyte donors. We did not find any studies evaluating male donors.

Key Points for Donors

Key findings for outcomes of sperm and oocyte donors included:
- For oocyte donors, observational studies suggest a lower incidence of OHSS with GnRH agonist trigger than with hCG trigger (low SOE). However, there was a lack of evidence on any long-term outcomes.
- There was a lack of evidence on any short or long-term outcomes for sperm donors.

Detailed Synthesis for Oocyte Donors

Included studies and their findings for the included studies are discussed below in terms of short- and long-term outcomes.

1. Short-Term Outcomes for Oocyte Donors

One included study was a cross-sectional survey published in 2009 (poor quality).271 The authors surveyed former oocyte donors who were registered on the Donor Sibling Registry, a U.S.-based worldwide registry designed to help donor-conceived individuals search for and contact their donor sibling. Of the 287 women with valid e-mail addresses who were invited via an email message to participate in the study, 155 (54%) completed the 25-item questionnaire. The mean age of respondents at the time of the survey was 35.8 years, and the number of years since first donation was 9.4 years (SD 5.2). Details of the stimulation protocols and oocyte retrieval methods were not reported. Of the 155 respondents, 47 (30.3%) reported some degree
of OHSS, 18 (11.6%) reported OHSS-related hospitalization and/or paracentesis. This study was rated as poor quality because of the high risk of bias due to the relatively low response rate and the self-selected nature of the study sample (insufficient SOE for all outcomes).

A randomized crossover trial conducted in Cyprus in 2008 compared outcomes associated with hCG versus GnRH agonist (leuprolide acetate) trigger among 50 oocyte donors.267 Both study protocols used recombinant FSH or hMG administered beginning on day 3 of each donor’s menstrual cycle for donor stimulation. Among the 44 donors who received both triggering agents in 2 consecutive cycles, 3 cases (6.8%) of mild-to-moderate OHSS were reported after administration of hCG. No cases of OHSS were reported after administration of the GnRH agonist leuprolide. The between-group difference in rates of OHSS was not statistically significant, but this trial may have been underpowered to detect a statistically significant effect for this clinical outcome. This study was rated as fair quality because of the small sample size and the exclusion from the analysis of patients who did not complete the trial (insufficient SOE).

Two publications reported retrospectively assessed outcomes observed among overlapping cohorts of oocyte donors treated at a private infertility clinic in Spain between 2001 and 2007.268,269 Among 1907 donors who collectively underwent 4052 stimulation cycles that reached oocyte retrieval, the stimulation protocol and triggering agents used were as follows: 1238 cycles with a GnRH agonist protocol and hCG trigger; 1295 cycles with a GnRH antagonist protocol and hCG trigger; and 1519 cycles with a GnRH antagonist protocol and a GnRH agonist trigger. The incidence of moderate or severe OHSS resulting from these three protocols were, respectively, 0.65 percent, 1.08 percent, and 0 percent (difference not statistically significant).268 The other publication from this cohort of patients reported an incidence of 13 cases of moderate or severe OHSS associated with 624 cycles of hCG triggering (2.1%) and no cases of moderate or severe OHSS associated with 547 cycles of triggering with a GnRH agonist.269 These two studies were rated as fair quality because of their retrospective design. Together they suggest a lower incidence of OHSS with GnRH agonist trigger than with hCG trigger (low SOE).

The fifth study was a retrospective analysis of all attempts at oocyte donations by anonymous and known directed donors at a medical center in the United States from 1991 to 2007.270 The charts of 587 donors (481 anonymous and 106 directed) who participated in 973 stimulation cycles and 886 retrievals were reviewed. The age of the donors ranged from 20 to 42 years. Of the 886 stimulation cycles, 12 (1.4%) were associated in mild or moderate OHSS that led to 2 outpatient office visits, and 4 (0.5%) were associated with moderate OHSS that led to 3 or 4 office visits. There was a single case (0.1%) of intraabdominal bleeding and 18 cases (2.0%) of other complications such as cysts, hematomas, urinary tract infections, yeast infections, or “vague symptoms.” This study was rated as fair quality because of its retrospective design and because details of the stimulation and retrieval protocols were not reported. In addition, the applicability of these findings is limited in that clinical practice have evolved since 1991 and how the protocol during the study period reflected these changes in practice is unclear (insufficient SOE).

2. Long-Term Outcomes for Oocyte Donors

The only included study with evidence on long-term outcomes was the 2009 cross-sectional survey referenced above by Kramer and colleagues (poor quality).271 The authors surveyed former oocyte donors who were registered on the Donor Sibling Registry, a U.S.-based worldwide registry designed to help donor-conceived individuals search for and contact their
donor sibling. Of the 287 women with valid e-mail addresses who were invited via an email message to participate in the study, 155 (54%) completed the 25-item questionnaire. The mean age of respondents at the time of the survey was 35.8 years, and the number of years since first donation was 9.4 years (SD 5.2). Details of the stimulation protocols and oocyte retrieval methods were not reported. Of the 155 respondents, and 41 (26.4%) reported infertility and/or menstrual changes since donation. Fifteen of these women (9.6% of the total sample) reported new infertility problems; of these, only 4 reported having become pregnant. This study was rated as poor quality because of the high risk of bias due to the relatively low response rate and the self-selected nature of the study sample (insufficient SOE for all outcomes)

**Strength of Evidence for Donors**

Table 28 summarizes the SOE for the incidence of OHSS with GnRH agonist trigger versus hCG trigger. All other short- and long-term outcomes had insufficient SOE or were not evaluated in the limited set of included studies.

**Table 28. Strength of evidence for major outcomes—KQ 6 (donor)**

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GnRH agonist (leuprolide acetate) vs. hCG trigger</td>
<td>Short term adverse effects of treatment: OHSS</td>
<td>2 Obs²⁶⁸,²⁶⁹ (3824)</td>
<td>Improvement: Lower incidence of OHSS with GnRH agonist trigger than with hCG trigger.</td>
<td>Low (Moderate study limitations, imprecise)</td>
</tr>
</tbody>
</table>

*Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.

Abbreviations: GnRH=gonadotropin-releasing hormone; hCG= human chorionic gonadotropin; Obs=observational study; OHSS=ovarian hyperstimulation syndrome

**Key Findings Across All Infertility Diagnoses**

Select key studies reported findings applicable across all KQs. Note these studies all adjusted for cause of infertility although did not report their findings for the specific causes of infertility and therefore are reported here rather than in the individual KQ sections. The findings from these studies are considered relevant to individuals with all included infertility diagnoses.

**Description of Included Studies Across All KQs**

We identified 26 articles²⁶,¹²⁵-¹²⁷,¹³⁶-¹³⁸,²⁷²-²⁹⁰ described in 21 studies that addressed outcomes after treatment for infertility and adjusted for cause of infertility.

Among the 21 studies, 3 were an RCTs and all were rated as good quality.²⁷⁵,²⁸⁵,²⁸⁸ The remaining 18 studies were observational studies; 16 of these were good quality,¹²⁵,¹³⁶,²⁷²,²⁷⁴,²⁷⁶-²⁷⁹,²⁸¹-²⁸³,²⁸⁶,²⁸⁷,²⁸⁹,²⁹⁰ and three were fair quality.²⁷³,²⁸⁰,²⁸⁴ Geographically, most of the 21 studies took place in the United States or the United Kingdom/Europe; 14 studies in the U.S.¹³⁶,²⁷²,²⁷⁴,²⁷⁶-²⁸⁴,²⁸⁶,²⁹⁰ and seven studies in the U.K./Europe.¹²⁵,²⁷³,²⁷⁵,²⁸⁵,²⁸⁷-²⁸⁹ Most of the studies (19 total) were conducted in or used data from fertility subspecialty clinics, while the remaining 2 did not specify the setting.²⁷²,²⁸² Eight studies reported support by government funding,¹¹⁶,²⁷²,²⁷⁴,²⁷⁶,²⁷⁷,²⁸²,²⁸⁴,²⁹⁰ two studies reported industry funding,²⁷⁵,²⁸⁵ three studies reported non-government, non-industry funding,²⁷³,²⁸⁰,²⁸⁷ three studies reported a combination of
Key Points for Any Infertility Diagnosis

Key points for patients who undergo IVF/ICSI include:

- Clomiphene or gonadotropins ever use was not associated with increased risk of maternal cancer (low SOE).
- Women who undergo IVF demonstrated an increased risk of ovarian neoplasms and colorectal malignancies (low SOE) compared to women who do not undergo IVF. There is no evidence of a difference in invasive ovarian cancers (low SOE).
- For children born after ART, ICSI may be associated with an increased risk of autism compared to IVF (low SOE).
- In the United States, live birth rates after IVF/ICSI are lower for African-Americans than for other racial/ethnic groups after adjusting for other prognostic factors (low SOE).
- Elective single-embryo transfer is associated with lower live birth rates but a significant reduction in multiple birth rates compared to multiple-embryo transfer (low SOE for both outcomes).
- There was no difference in the odds of low birth weight between ICSI versus conventional IVF cycles (low SOE). However, among couples undergoing ART with a singleton pregnancy, frozen embryo transfers result in a higher average birthweight, with a subsequent reduction in the incidence of low birthweight and an increase in the incidence of macrosomia (low SOE).

Long-Term Outcomes: Maternal

Cancer

Two studies evaluated long-term risk for cancer in infertile women (by diagnosis) who were treated with clomiphene citrate, gonadotropins, or IVF and reported results in multiple publications, as follows:

- Brinton, 2015: Primary report\textsuperscript{136} and three companion papers\textsuperscript{26,137,138}
- Spaan, 2015: Primary report\textsuperscript{125} and a companion paper\textsuperscript{126}

One study compared the long-term risk for cancer in infertile women treated with clomiphene citrate versus gonadotropins, with risks for different cancers reported in four separate papers.\textsuperscript{26,136-138} This study was a prospective cohort of 12,193 women who sought treatment for infertility between 1965 and 1988 at five reproductive endocrinology practices in the United States. The analytic cohort comprised 9892 women successfully traced, with a median follow-up of approximately 30 years.

Tables 29 and 30 highlight the risk of cancer in relation to use of clomiphene citrate or gonadotropins in women by infertility diagnosis.
### Table 29. Cancer risk (95% CI) for ever use of clomiphene citrate by infertility diagnosis

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Endometriosis (N=2196)</th>
<th>Tubal or Peritoneal Factor (N=3496)</th>
<th>Male Factor (N=2218)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ovarian</td>
<td>1.01 (0.43 to 2.36)</td>
<td>0.98 (0.46 to 2.07)</td>
<td>1.18 (0.52 to 2.68)</td>
</tr>
<tr>
<td>Endometrial</td>
<td>1.79 (0.76 to 4.23)</td>
<td>1.14 (0.59 to 2.21)</td>
<td>1.55 (0.82 to 2.96)</td>
</tr>
<tr>
<td>Breast</td>
<td>1.23 (0.92 to 1.65)</td>
<td>1.12 (0.87 to 1.45)</td>
<td>1.25 (0.92 to 1.69)</td>
</tr>
<tr>
<td>Colon</td>
<td>0.53 (0.22 to 1.27)</td>
<td>0.66 (0.31 to 1.40)</td>
<td>1.14 (0.42 to 3.06)</td>
</tr>
<tr>
<td>Lung</td>
<td>1.68 (0.71 to 3.99)</td>
<td>1.78 (0.90 to 3.54)</td>
<td>1.35 (0.51 to 3.61)</td>
</tr>
<tr>
<td>Thyroid</td>
<td>0.87 (0.28 to 2.67)</td>
<td>1.08 (0.40 to 2.88)</td>
<td>2.46 (0.64 to 9.49)</td>
</tr>
<tr>
<td>Melanoma</td>
<td>1.10 (0.42 to 2.89)</td>
<td>0.98 (0.36 to 2.73)</td>
<td>1.10 (0.27 to 4.46)</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; N=number of patients

### Table 30. Cancer risk (95% CI) for ever use of gonadotropins by infertility diagnosis

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Endometriosis (N=2196)</th>
<th>Tubal or Peritoneal Factor (N=3496)</th>
<th>Male Factor (N=2218)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometriosis</td>
<td>1.15 (0.72 to 1.83)</td>
<td>1.16 (0.77 to 1.74)</td>
<td>1.29 (0.79 to 2.11)</td>
</tr>
<tr>
<td>Breast</td>
<td>2.74 (0.96 to 7.85)</td>
<td>0.94 (0.32 to 2.70)</td>
<td>0.74 (0.22 to 2.43)</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; N=number of patients

While the findings from this cohort are suggestive that, in couples with infertility due to endometriosis, tubal factors, or male factors, that women treated with clomiphene citrate or gonadotropins are not at markedly increased risk for any common cancer, the size of the cohort was not adequate to detect modest increases in risk. In addition, the size of the cohort precluded conducting more detailed analyses in relation to the number of cycles received or age at treatment (low SOE).

Cancer risk in relation to ART treatment was examined in two good-quality observational cohort studies, described in four papers. Results from a study conducted in the Netherlands were described in papers examining melanoma risk, ovarian neoplasms, and colorectal cancers. In this study, the underlying cause of infertility was adjusted for along with other potential confounders. The cohort was comprised of 19,158 women who received IVF treatment and 6006 infertile women who did not receive IVF treatment after a median follow-up of 15 years. IVF was not significantly associated with melanoma risk (standardized incidence ratio [(SIR] 1.27; 95% CI, 0.75 to 2.15) when controlling for cause of infertility. IVF was associated with a statistically significant increased risk of all ovarian malignancies (HR 2.05; 95% CI, 1.10 to 3.82). This risk was increased most notably for borderline ovarian tumors (HR 6.38; 95% CI, 2.05 to 19.84) as compared to the increase seen in invasive ovarian cancer (HR 1.14; 95% CI, 0.54 to 2.41) (low SOE). For colorectal cancers, IVF was not associated with a significant risk of colorectal cancer compared to the general population (SIR 1.00; 95% CI, 0.80 to 1.23), but was increased compared to infertile patients who did not receive IVF (HR 1.80, 95% CI, 1.10-2.94). Interestingly, this may be due to a significantly lower risk for colorectal cancers in this group compared to the general population (SIR 0.58, 95% CI, 0.36-0.88), which the authors speculated may have been due to a “healthy female effect”, but could also have been a chance finding because of a small number of cases in the non-IVF group.

The applicability of the findings may be limited since the study was conducted in the Netherlands and the IVF treatments were received in the period between 1983 and 1995. With the evolution of IVF treatment over time, findings from this study may not reflect the risk for women currently undergoing IVF.

The second study, conducted in the UK, examined ovarian, breast and corpus uteri cancers in a cohort of 255,786 women who underwent ART and were followed for an average of 8.8 years. SIRs were calculated comparing the observed incidence rates to expected rates based on
national cancer incidence rates in women and are summarized in Table 31. Breast cancer risk (invasive or \textit{in situ}) was not significantly increased in any of the female infertility diagnoses categories, and was significantly lower among those with a male factor diagnosis. Both invasive and borderline ovarian cancers were significantly increased among women with a diagnosis of endometriosis or tubal disease. Corpus uteri cancer risk was significantly higher only among women with a diagnosis of ovulatory problems.

### Table 31. Cancer risk (SIR 95% CI) among women who underwent assisted reproduction by infertility diagnosis

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Endometriosis (181,279 Person-Years)</th>
<th>Tubal Disease (710,522 Person-Years)</th>
<th>Ovulatory Problems (311,523 Person-Years)</th>
<th>Male Factor (757,063 Person-Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invasive breast cancer</td>
<td>0.95 (0.82 to 1.10)</td>
<td>0.94 (0.87 to 1.01)</td>
<td>0.91 (0.81 to 1.02)</td>
<td>0.89 (0.83 to 0.96)</td>
</tr>
<tr>
<td>In situ breast cancer</td>
<td>1.25 (0.81 to 1.83)</td>
<td>1.11 (0.89 to 1.36)</td>
<td>1.05 (0.75 to 1.42)</td>
<td>1.18 (0.95 to 1.44)</td>
</tr>
<tr>
<td>Invasive ovarian tumors</td>
<td>2.47 (1.75 to 3.39)</td>
<td>1.71 (1.40 to 2.08)</td>
<td>1.16 (0.80 to 1.63)</td>
<td>1.09 (0.84 to 1.39)</td>
</tr>
<tr>
<td>Borderline ovarian tumors</td>
<td>2.03 (1.18 to 3.25)</td>
<td>1.62 (1.21 to 2.12)</td>
<td>1.52 (0.96 to 2.31)</td>
<td>0.96 (0.66 to 1.35)</td>
</tr>
<tr>
<td>Corpus uteri cancer</td>
<td>0.75 (0.35 to 1.43)</td>
<td>1.23 (0.93 to 1.58)</td>
<td>1.59 (1.13 to 2.17)</td>
<td>0.91 (0.65 to 1.24)</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; N=number of patients; SIR=standardized incidence ratio

### Short-Term Outcomes: Child

#### Birthweight

A good-quality observational study reported examined birth outcomes among all live births in Massachusetts (2004 through 2008) that linked to ART cycles in the Society for Assisted Reproductive Technology Clinic Online Reporting System (SART CORS) and the Pregnancy to Early Life Longitudinal (PELL) data system.\textsuperscript{274} There was no difference in low birth weight by assisted hatching, donor/autologous source of oocytes/semen, number of embryos or ICSI (low SOE).\textsuperscript{274}

In a good-quality NASS comparing fresh and frozen embryo transfer, singleton infants both after frozen/thawed transfers had an average birthweight 142.3 grams heavier than those born after fresh transfers. In terms of clinically relevant categories, this translated into a significantly lower risk for low birthweight after frozen/thawed transfer (aRR 0.52, 95% CI 0.48 to 0.56), and a significantly increased for macrosomia (aRR 1.70, 95% CI 1.64 to 1.76).\textsuperscript{284}

Finally, analyses of the SART CORS database compared outcomes for ICSI versus conventional IVF.\textsuperscript{185} Keyhan, et al. reported on low birth weight (<2500 g) in relation to ICSI versus conventional IVF using propensity score matching.\textsuperscript{185} The propensity score matching, which included infertility diagnosis as one of the factors used to calculate propensity scores, resulted in 12,364 ICSI cycles and 12,364 IVF cycles. In the ICSI group, 9.4% of the births were low birth weight compared to 9.7% in the IVF group (OR 0.92, 95% CI 0.78-1.10, p=0.35) (low SOE).

#### Congenital Anomalies

A good-quality observational study\textsuperscript{277} using the States Monitoring ART (SMART) Collaborative and linked ART surveillance, birth certificates, and birth defects registry data for Florida, Massachusetts, and Michigan from 2000 to 2010 (n=64,861). The objective of this study
was to examine the prevalence of birth defects among liveborn infants conceived with and without ART and to evaluate risks associated with certain ART procedures among ART-conceived infants. Overall, the prevalence ratio was significantly higher for ART versus non-ART birth (adjusted risk ratio 1.28, 95% CI 1.15 to 1.42, p<0.001) (low SOE).

Long-Term Outcomes: Child

Neurodevelopmental Outcomes

Risk of autism through age 5 was examined in one good-quality observational study involving 42,383 children conceived with ART in California between 1997 and 2006. Analyses did not examine associations with type of embryo fertilization (ICSI or conventional IVF) stratified by cause of infertility, but did control for infertility diagnosis in multivariable analyses. Among ART-conceived infants, use of ICSI was associated with a higher incidence of autism in both singleton births (HR 1.65; 95% CI, 1.08 to 2.52) and multiple births (HR 1.71; 95% CI, 1.10 to 2.66) in multivariable analyses that controlled for infertility diagnosis (low SOE). Results of ICSI versus IVF were not described for specific causes of infertility, although male factor infertility was the most common diagnosis in the cohort.

One fair-quality observational study compared outcomes between ART and no intervention/expectant management. Risk of neurological dysfunction at 2 years was assessed in a fair-quality observational study examining children conceived with IVF (n=122), conceived naturally to subfertile parents (n=87), or born to parents without fertility problems (n=101). Outcomes reported were simple or complex minor neurological dysfunction. Results were stratified by underlying cause of infertility. None of the specific causes of infertility were related to the rates of neurological dysfunction (insufficient SOE).

Subgroups of Interest (With Any Infertility Diagnosis)

Race

One good-quality observational study assessed the association between race and IVF outcomes. The live birth rate per patient was 34.7% among whites, 19.8% among blacks, 33.3% among South Asians, 31.3% among those who reported mixed race, 28.4% among those who reported “other,” and 36.1% among those who reported “unknown” (p<0.001 for black vs. white comparison; p>0.05 for South Asian vs. white) (low SOE).

Number of Embryos Transferred

A fair-quality observational study using SART CORS data assessed 69,028 cycles to determine factors associated with elective single-embryo transfer (eSET) and controlled for infertility diagnosis. Findings from the study are summarized in Table 32, demonstrating increased live birth rate per cycle with 2 embryo transfer as compared to single-embryo transfer (low SOE) and that multiple live birth rates are significantly higher with a 2-embryo transfer than a single-embryo transfer, but do not increase further with 3- or 4-embryo transfers (low SOE).
Table 32. Birth rates by number of embryos transferred

<table>
<thead>
<tr>
<th># Embryos Transferred</th>
<th>1 Embryo Transferred</th>
<th>2 Embryos Transferred</th>
<th>3 Embryos Transferred</th>
<th>4+ Embryos Transferred</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singleton birth</td>
<td>1,302/3,037 (42.9%)</td>
<td>13,779/42,396 (32.5%)</td>
<td>4,632/17,480 (26.5%)</td>
<td>1,424/6,115 (23.3%)</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Multiple birth</td>
<td>15/3,037 (0.5%)</td>
<td>8,055/42,396 (19.0%)</td>
<td>3,094/17,480 (17.7%)</td>
<td>923/6,115 (15.1%)</td>
<td>p&lt;0.0001</td>
</tr>
<tr>
<td>Total births per cycle</td>
<td>1,318/3,037 (43.4%)</td>
<td>21,834/42,396 (51.5%)</td>
<td>7,726/17,480 (44.2%)</td>
<td>2,348/6,115 (38.4%)</td>
<td>p&lt;0.0001</td>
</tr>
</tbody>
</table>

A good-quality study from NASS assessed correlation between infertility clinic eSET rates and pregnancy outcomes—there was a significant linear decrease in multiple birth rates with increasing eSET rates with no significant differences in clinic-level live birth rates for women younger than 38 years.283

**Number of Oocytes Retrieved**

A good-quality observational study used data from the Swedish National Quality Registry of Assisted Reproduction and Medical Birth/IVF Registry assessed the association between the number of oocytes retrieved and obstetric and neonatal outcomes among 27,359 women who delivered singleton babies after IVF.287 In multivariable models that adjusted for cause of infertility as well as multiple maternal characteristics, the number of oocytes retrieved (continuous) was not significantly associated with pre-term birth (OR 1.002, 95% CI 0.994-1.011), peri/neonatal death (OR 1.008, 95% CI 0.975-1.043), or major birth defects (OR 1.009, 95% CI 0.998-1.001).

**Fresh Versus Frozen IVF Cycles**

Five observational studies evaluated outcomes using fresh or frozen IVF cycles. A good-quality observational study using SART CORS data compared pregnancy outcomes in 509,938 IVF cycles, based on the transfer of a fresh or frozen blastocyst or non-blastocyst.286 Table 33 summarizes findings from multivariable models that controlled for infertility diagnosis and shows that the best outcomes in terms of live births and first trimester pregnancy loss for fresh blastocyst transfer, and the worst outcomes for frozen non-blastocyst transfers. The risk for ectopic pregnancies was lower for frozen transfers, whether blastocyst or non-blastocyst.

Table 33. Outcomes with fresh and frozen IVF cycles

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Live Birth OR (95% CI)</th>
<th>First-Trimester Pregnancy Loss OR (95% CI)</th>
<th>Ectopic/Heterotopic Pregnancy OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fresh blastocyst</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Fresh non-blastocyst</td>
<td>0.82 (0.80 to 0.83)</td>
<td>1.15 (1.12 to 1.18)</td>
<td>1.04 (0.98 to 1.11)</td>
</tr>
<tr>
<td>Frozen blastocyst</td>
<td>0.73 (0.72 to 0.75)</td>
<td>1.29 (1.25 to 1.32)</td>
<td>0.48 (0.43 to 0.53)</td>
</tr>
<tr>
<td>Frozen non-blastocyst</td>
<td>0.64 (0.62 to 0.65)</td>
<td>1.34 (1.30 to 1.38)</td>
<td>0.67 (0.60 to 0.73)</td>
</tr>
</tbody>
</table>

Abbreviations: CI=confidence interval; OR=odds ratio

Another good-quality SART CORS study found no difference in live birth or miscarriage rate for fresh or frozen oocytes in autologous cycles, but a significantly lower rate of live birth with frozen oocytes in donor cycles (adjusted Risk Ratio [aRR] 0.87, 95% CI, 0.80 to 0.95), with no difference in miscarriage rates.281

A good-quality observational study examined birth outcomes among all live births in Massachusetts (2004 through 2008) that linked to ART cycles in the Society for Assisted Reproductive Technology Clinic Online Reporting System (SART CORS) and the Pregnancy to Early Life Longitudinal (PELL) data system274 and found lower birthweight with thawed
embryos compared to fresh (adjusted OR and 95% CI 0.79, 0.65 to 0.96). There was no difference in low birth weight by assisted hatching, donor/autologous source of oocytes/semen, number of embryos or ICSI.274

Finally, within frozen cycles, another good-quality SART CORS study compared the live birth rate in first-cycle frozen embryo transfers with and without assisted hatching using propensity score matching.282 The propensity score matching, which included infertility diagnosis as one of the factors used to calculated propensity scores, resulted in 70,738 assisted hatching cycles and 80,795 cycles without assisted hatching. The live birth rate was significantly lower in the assisted hatching cohort compared to the no assisted hatching cohort (34.2% versus 35.4%, p<0.001).

Blastocyst or Cleavage Stage Embryo Transfer

A good-quality observational study used CDC’s National ART Surveillance System (NASS) to evaluate differences in birthweight in 124,154 infants born after embryo transfer at the blastocyst stage or the cleavage stage.290 The results of the multivariable linear model that controlled for infertility diagnosis, maternal demographics and pregnancy characteristics showed that infants born after blastocyst transfer had birth weights that were slightly but statistically significantly higher than those with cleavage stage transfer (5.73 grams, p=0.040). The difference was larger for single-embryo transfer (19.26 grams, p=0.008) than for double embryo transfer (4.03 grams, p=0.245). Embryo transfer type was not associated with low birth weight (RR 1.0, 95% CI 0.96-1.04).

GnRH Suppression Protocols

A good-quality observational study using SART data compared ectopic pregnancy rates outcome of ectopic pregnancy in patients who used different GnRH analog protocols for ovarian hyperstimulation prior to IVF in fresh autologous cycles.278 This study compared this outcome in women who received one of 3 treatments: (1) a luteal phase GnRH agonist protocol; 2) a GnRH agonist flare protocol, begun during the follicular phase; or 3) a GnRH antagonist suppression protocol, begun in the mid-follicular phase. In the overall population, as compared to luteal GnRH agonist cycles in which ectopic pregnancies occurred in 1.6% of the cycles, a higher risk for ectopic pregnancy was reported for both GnRH antagonist cycles (2.4%, OR 1.52 (95% CI 1.39 to 1.65)) and GnRH agonist flare cycles (2.1%, OR 1.25 (95% CI 1.09 to 1.44)). In multivariate models, among women with unexplained infertility, compared to women receiving the luteal phase GnRH agonist protocol, those receiving the GnRH agonist flare protocol had a lower risk of ectopic pregnancy with an adjusted OR (95% CI) of 0.60 (0.37 to 0.96), P=0.03. Similarly, in multivariate models, among women with unexplained infertility, compared to women receiving the GnRH antagonist protocol, those receiving the GnRH agonist flare protocol had a lower risk of ectopic pregnancy with an adjusted OR (95% CI) of 0.50 (0.32 to 0.80), P=0.004.278 There were no significant differences across protocols for PCOS, endometriosis, or tubal factor diagnoses.

Preimplantation Genetic Diagnosis (PGD) Cycles With ART Versus Non-PGD Cycles

In a good-quality observational study using CDC’s web-based NASS database (which includes data from clinics participating in SART CORS, as well as a smaller number of non-SART participating clinics who report directly to CDC), for women under age 35, the odds of
live birth per transfer was lower for all types of PGD cycles when compared with non-PGD cycles. Among live-birth deliveries, in models adjusting for cause of infertility, the adjusted OR (95% CI) for low birth weight among PGD-Genetic cycles was 0.73 (0.54 to 0.98) as compared to those resulting from non-PGD cycles. In contrast, the adjusted OR (95% CI) for low birth weight among PGD-Aneuploidy cycles was 1.25 (1.01 to 1.54) compared with non-PGD cycles. Among live births, the adjusted OR (95% CI) for multiple birth in PGD-Other cycles was 0.76 (0.60 to 0.97) compared with non-PGD cycles. Results for women 35-37 and >37 years of age for low birth weight among PGD-Genetic cycles as compared to those resulting from non-PGD cycles or for PGD-Aneuploidy cycles compared with non-PGD cycles were not statistically significantly different. Results for women 35-37 and >37 years of age for multiple births among PGD-Genetic cycles as compared to those resulting from non-PGD cycles or for PGD-Aneuploidy cycles compared with non-PGD cycles were also not statistically significantly different. 

In another good-quality study from SART CORS, PGD was associated with a lower chance of live birth with donor oocyte cycles compared to non-PGD cycles (OR 0.65, 95% CI, 0.53 to 0.80).

Insurance Coverage

A cohort study of the SART CORS database examined pregnancy in States with mandated insurance of IVF coverage as compared with States without mandated IVF coverage. In models adjusted for cause of infertility, there was no significant difference in the odds of live birth in states with mandated insurance IVF coverage as compared with states with non-mandated IVF coverage (aOR 1.02, 95% CI 0.97 to 1.07). The odds of multiple birth or higher-order multiple birth were lower in States with mandated insurance IVF coverage as compared with States with non-mandated IVF coverage. The adjusted odds ratio (95% CI) of multiple birth or higher order multiple-birth in States with mandated IVF coverage was 0.87 (0.80 to 0.94) and 0.74 (0.53 to 1.03), respectively, as compared with States without mandated IVF coverage. There was no statistically significant difference in the odds of low birth weight in States with mandated IVF coverage as compared with States without mandated IVF coverage (aOR 0.95, 95% CI 0.88 to 1.03).

Strength of Evidence (With Any Infertility Diagnosis)

Table 34 summarizes the SOE findings which span all infertility diagnoses.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Outcome</th>
<th>Study Design (Sample Size)</th>
<th>Conclusion</th>
<th>Strength of Evidence (Rationale)⁹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clomiphene citrate and gonadotropin</td>
<td>Long-term outcomes: Maternal cancer</td>
<td>1 Obs¹⁵⁶ (9892 patients)</td>
<td>No difference. Ever use of clomiphene citrate was not statistically significantly associated with maternal ovarian, breast, endometrial, lung, thyroid, colon, or melanoma cancer. Gonadotropin use was not associated with increased risk for breast or endometrial cancer</td>
<td>Low (Size of cohort not sufficient to detect modest increases in risk)</td>
</tr>
<tr>
<td>Comparison</td>
<td>Outcome</td>
<td>Study Design (Sample Size)</td>
<td>Conclusion</td>
<td>Strength of Evidence (Rationale)*</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>----------------------------</td>
<td>------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>ART: IVF</td>
<td>Live birth (by race)</td>
<td>1 Obs $^{192}$ (13,473 cycles)</td>
<td>Greater disparity. Lower live birth rate for blacks as compared to white (p&lt;0.001)</td>
<td>Low (Imprecise, 1 study)</td>
</tr>
<tr>
<td></td>
<td>Live birth (by number of embryos transferred)</td>
<td>1 Obs $^{23}$ (69,028 cycles)</td>
<td>Improvement. Increased live birth rate per cycle with 2 embryo transfer as compared to single-embryo transfer</td>
<td>Low (Imprecise, findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Pregnancy complications: Multiple births (by number of embryos transferred)</td>
<td>1 Obs $^{23}$ (69,028 cycles)</td>
<td>Greater risk. Multiple live birth rates are significantly higher with a 2-embryo transfer than a single-embryo transfer, but do not increase further with 3- or 4-embryo transfers</td>
<td>Low (Imprecise, findings with moderate study limitations)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birthweight</td>
<td>1 Obs $^{24}$ (8,948)</td>
<td>No difference: No significant difference in rates of low birthweight using ART by assisted hatching, source of oocytes/semen, number of embryos or ICSI</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Congenital Anomalies</td>
<td>1 Obs $^{37}$ (64,861)</td>
<td>Greater risk. Risk of birth defects was greater in infants conceived using ART (adjusted risk ratio 1.28, 95% CI 1.15 to 1.42)</td>
<td>Low (1 study)</td>
</tr>
<tr>
<td></td>
<td>Long-term outcomes: Child (Autism)</td>
<td>1 Obs $^{72}$ (42,383)</td>
<td>Greater risk. Risk of autism was greater in children conceived with ART with ICSI as compared to ART without ICSI (HR 1.65, p&lt;0.05)</td>
<td>Low (Imprecise)</td>
</tr>
<tr>
<td></td>
<td>Long-term outcomes: Child (neurological)</td>
<td>1 Obs $^{73}$ (310 patients)</td>
<td>Inconclusive. SOE was insufficient given imprecise evidence from 1 small observational study with moderate risk of bias</td>
<td>Insufficient (imprecise findings with moderate study limitations, small study)</td>
</tr>
<tr>
<td></td>
<td>Long-term outcomes: Maternal (cancer)</td>
<td>2 Obs $^{25,26}$ (280,950)</td>
<td>Greater risk. IVF was associated with a statistically significant increased risk of all ovarian neoplasms (HR 2.05; 95% CI, 1.20 to 3.82) and borderline ovarian tumors (HR 6.28; 95% CI, 2.05 to 19.84), and colorectal cancer (HR 1.80, 95% CI 1.10 to 2.94)</td>
<td>Low (Imprecise, older study)</td>
</tr>
<tr>
<td></td>
<td>Neonatal outcomes: Birthweight</td>
<td>1 Obs $^{85}$ (90,401 cycles)</td>
<td>No difference: No significant difference in the odds of low birth weight between ICSI versus conventional-IVF cycles</td>
<td>Low (1 study with moderate study limitations)</td>
</tr>
</tbody>
</table>

*Criteria for downgrading strength of evidence is described as Rationale; when these criteria are insufficient for understanding the final strength of evidence, additional explanation is provided.
Publication Bias

As part of the Evidence-based Practice Center (EPC) Program, AHRQ sought to assess whether information from ClinicalTrials.gov would impact the conclusions of five ongoing systematic reviews. This infertility systematic review was part of this methods project (Augmenting Systematic Reviews with Information from ClinicalTrials.gov to Increase Transparency and Reduce Bias).

For this purpose, we searched the ClinicalTrials.gov registry of clinical studies to ascertain publication bias by identifying studies that have been completed but are as yet unpublished. Our search yielded 354 records of completed trials about treatments for infertility for screening (see Appendix A for our search strategy and Appendix H for details on our findings). Initial manual review identified 94 of these records as potentially relevant; subsequent review by a topic expert reduced this number to 66. Of these 66 records, we were not able to identify publications for 12 studies that had expected completion dates 3 years or more prior to our search. During the search update period we again looked for publications covering these 12 studies. No new publications were found.

Of these 12 trials with unpublished results, two were considered potentially relevant to KQ 1, 9 potentially relevant to KQ 3, and 1 potentially relevant to KQ 5. Implications for KQs 1, 2, and 4 are discussed in more detail in Appendix H as part of the augmenting transparency methods project.

For this current systematic review, the two trials identified as potentially relevant to KQ 1 had a combined sample size of 340 patients and were completed more than 3 years ago. These two “missing” trials are unlikely to have had a meaningful impact on our review’s results especially given the presence of two systematic reviews for women with infertility from PCOS.

For KQ 3, there were nine potentially relevant unpublished studies where the underlying diagnosis of infertility was not listed. It is not clear whether these trials are specifically relevant to patients with infertility for unknown reasons as is required for inclusion in KQ 3 since the trials may focus on patients with identified infertility diagnoses. This diagnosis specification however is not available from the missing trials. The nine studies differed in terms of treatments evaluated and outcomes assessed. All studies were fairly small studies (planned enrollment varying between 60 and 242 individuals) and most of them focused on pregnancy rates rather than live births though could potentially be relevant based on information about miscarriages, ovarian hyperstimulation syndrome (OHSS), or other adverse events.

For KQ 5, there was one unpublished potentially relevant study which sought to determine whether couples with male factor infertility, specifically with elevated sperm DNA damage undergoing in vitro fertilization (IVF)/intra-cytoplasmic sperm injection (ICSI), should use testicular sperm extraction to improve their reproductive outcomes. Although this study did have live birth listed as the primary outcome, it had an enrollment of 25 males and so would not likely change any of our findings.

In summary, because of the relatively low number of unpublished studies identified through our ClinicalTrials.gov registry analysis as compared to our included set of studies, we do not believe these findings indicate significant publication bias in the evidence base that would impact our overall conclusions.
Discussion

In this Comparative Effectiveness Review, we reviewed 151 studies described in 161 publications that directly compared infertility management strategies in couples with infertility due to polycystic ovary syndrome (PCOS; Key Question [KQ] 1) or endometriosis (KQ 2); unexplained infertility (KQ 3); tubal and peritoneal factor infertility (KQ 4); and male factor infertility (KQ 5). We also explored the comparative safety and effectiveness of management strategies for donors in infertility (KQ 6). Although the ultimate goal with any infertility management strategy is to improve live birth rates of healthy infants to a healthy couple, many studies initially identified in our review only reported on pregnancy rates or focused on other short-term outcomes and did not differentiate by the underlying causes of infertility. Our findings are based on those 151 studies which evaluated the comparative effectiveness of infertility management strategies in couples with a known cause of infertility (including unexplained infertility) and which evaluated the outcome of live birth or another long-term outcome.

Key Findings and Strength of Evidence

For women with infertility associated with PCOS, there was moderate strength of evidence (SOE) that letrozole compared to clomiphene results in higher live birth rates while reducing multiple births, with no difference in ectopic pregnancy or miscarriage (moderate SOE), or low birthweight and time to pregnancy (low SOE). There was moderate SOE clomiphene does not result in higher live birth rates compared to metformin, and low SOE for lack of differences in multiple birth, ectopic pregnancy, or time to pregnancy. Live birth rates are not different comparing laparoscopic ovarian drilling (LOD) with oral agents (moderate SOE).

For couples with endometriosis as the primary cause, there was insufficient evidence for specific comparisons/outcomes.

For couples with unexplained infertility, there is no difference between the oral agents of letrozole and anastrozole for the outcome of ectopic pregnancy (low SOE) but evidence is insufficient for other outcomes of interest. There is also no difference between differing adjunct treatments used in combination with oral agents and intrauterine insemination (IUI) for the outcomes of live birth, miscarriage and ovarian hyperstimulation syndrome (OHSS) (low SOE for all outcomes). Time to pregnancy was shorter with immediate in vitro fertilization (IVF) compared to strategies starting with clomiphene and IUI or gonadotropins and IUI followed by IVF if necessary (moderate SOE).

For couples with male factor infertility, live birth rate (moderate SOE) and miscarriage (low SOE) did not differ between intracytoplasmic sperm injection (ICSI) and intracytoplasmic morphological sperm injection (IMSI), a finding of limited applicability given the lack of clinical use of IMSI.

For oocyte donors, studies suggest a lower incidence of ovarian hyperstimulation syndrome (OHSS) with gonadotropin-releasing hormone (GnRH) agonist trigger than with human chorionic gonadotropin (hCG) trigger (low SOE). However, there was a lack of evidence on any long-term outcomes.

Findings applicable across all indications for infertility for couples undergoing assisted reproductive technology (ART) included: lower live birth rates for African-Americans compared to other racial/ethnic groups (low SOE); slightly lower live birth rates but significant reductions in multiple birth rates with elective single-embryo transfer compared to multiple-embryo transfer (low SOE); no increase in most maternal cancers after ART treatment after adjustment for
infertility in general or specific causes (low SOE), and, for children born after ART, a possible increased risk of neurodevelopmental disorders after ICSI compared to IVF (low SOE) but no evidence of an increased risk of Type I diabetes.

Findings in Relation to What Is Already Known

The 2008 Agency for Healthcare Research and Quality (AHRQ) Evidence Report on “Effectiveness of ART” found that approximately 80 percent of the 478 included studies were performed outside the United States, and that the majority of RCTs did not report delivery rates and obstetric outcomes. In that review, most studies did not have sufficient power to detect clinically meaningful differences in live birth rates, and had still lower power to detect differences in less frequent outcomes such as multiple births and complications. In addition, the previous report focused on outcomes of specific treatments (ovulation induction, superovulation, and IVF/ICSI) rather than a wider range of potential treatments, and infertility diagnosis was considered as subgroup analyses, rather than the primary basis for comparing treatments.

Methods for evidence synthesis, in particular for rating strength of evidence, have also been revised since that report. Although an increasing number of studies are using live birth rate as the primary outcome, the majority of the literature, particularly randomized trials, is still based on pregnancy or ongoing pregnancy. Lack of precision for comparative estimates of rates for less common but important outcomes, such as complications, continues to be a major limitation.

To put the findings of our present systematic review in context, Tables 35 through 41 provide a comparison of the findings of this review, by KQ, with two major sets of guidelines/recommendations—those of the National Institute for Health and Care Excellence (NICE) in the United Kingdom (UK), and those of the American Society for Reproductive Medicine (ASRM). Recommendations for NICE and ASRM are presented verbatim. For the NICE guidelines, recommendations are taken from the 2013 document, except as noted in the tables. References to specific recommendations by ASRM are provided within the tables. For simplicity, we do not present the results of relevant Cochrane reviews here; the NICE guidance relied heavily on available Cochrane reviews, and the specific recommendations reflect judgments about overall strength of evidence. Sections or statements shaded in gray represent outcomes where our review’s findings were different from those of the major guidelines/recommendations. Note that since our review focused on studies that reported live births—and not just pregnancies—several of the comparisons could not be made directly.

These tables demonstrate that in general, findings of our present review were concordant with the guidelines, with differences primarily attributable to differences in inclusion/exclusion criteria (particularly for publication dates and primary outcome of live birth vs. pregnancy).

For women with PCOS (Table 35), both NICE and ASRM support use of clomiphene citrate alone as first-line therapy, with the NICE guidance recommending ultrasound monitoring for dose adjustment to minimize risk of multiple pregnancy, followed by combination therapy with metformin or gonadotropins for women who do not conceive after a 3-6 month course of clomiphene alone. Both our review and NICE suggest letrozole may be superior to clomiphene as first line therapy, and that pretreatment with metformin may improve outcomes in women with PCOS being treated with gonadotropins.

For women with endometriosis (Table 36), ASRM concluded that evidence for surgical treatment of women with mild to moderate endometriosis was insufficient to recommend treatment, while the NICE guidance suggests some benefit, and our review was inconclusive. For
those patients going directly to ART, surgical treatment of endometriosis, including endometrioma, prior to ART does not improve outcomes.

For women with unexplained infertility (Table 37), NICE recommends against use of oral agents entirely, while ASRM suggests clomiphene plus IUI may improve cycle fecundity compared to expectant management; our review found insufficient evidence. Based on our review, immediate IVF results in higher live birth rates and shorter time to pregnancy in women aged 38-42 compared with a trial of clomiphene and IUI or gonadotropins and IUI, with most live births ultimately resulting from IVF.

For women with suspected tubal factor infertility (Table 38), both NICE and ASRM recommend imaging for diagnosis (which is outside the scope of our review), although, when ART is readily available and affordable, proceeding directly to ART without a definitive diagnosis of tubal disease may be more efficient.

For male factor infertility (Table 39), our review found no relevant findings compared to the recommendations, primarily because of limited data on live birth outcomes.

For both male and female donors (Table 40), both NICE and ASRM recommend psychological evaluation and counseling, including, for females, the short term risks of ovarian stimulation and oocyte collection; our review found evidence on outcomes was limited only to the known short-term risks of these procedures, with no evidence on potential longer term risks.

For long-term outcomes in women and children after infertility treatment (Table 41), our review found limited or inconsistent evidence. Risks of adverse longer term maternal cancer outcomes were generally not increased after adjustment for the risk associated with infertility itself. ICSI however may be associated with an increased risk of neurodevelopmental disorders in children compared to those conceived through IVF. The NICE guidance was generally consistent with this assessment, and recommended that patients should be informed that any absolute risk was low, while there was still uncertainty about longer-term outcomes.
<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selective estrogen receptor modulators</td>
<td>Clomiphene citrate does not result in higher live birth rates compared with metformin (low SOE). Differences are also not found in the rates of miscarriage, multiple birth, ectopic pregnancy, or time to pregnancy (low SOE for all outcomes)</td>
<td>For women with WHO Group II ovulation disorders (PCOS), offer as initial treatment clomiphene, metformin, or a combination, considering potential adverse effects, ease and mode of use, the woman's BMI, and monitoring needed&lt;br&gt;For women who are taking clomiphene citrate, offer ultrasound monitoring during at least the first cycle of treatment to ensure that they are taking a dose that minimizes the risk of multiple pregnancy.&lt;br&gt;For women who are taking clomiphene citrate, do not continue treatment for longer than 6 months. (Monitoring as part of treatment was not included in our review.)</td>
<td>Clomiphene citrate is an effective first-line treatment for the majority of women with anovulatory infertility. Failure to conceive after 3 to 4 successful CC-induced ovulation cycles is indication for further evaluation to exclude other contributing causes of infertility, particularly in women &gt;35 years of age.</td>
</tr>
<tr>
<td>Insulin sensitizers</td>
<td>Clomiphene citrate does not result in higher live birth rates compared with metformin (low SOE). Differences are also not found in the rates of miscarriage, multiple birth, ectopic pregnancy, or time to pregnancy (low SOE for all outcomes)&lt;br&gt;The combination of metformin and clomiphene does not significantly improve live birth rates in women with higher BMI, but precision is limited.</td>
<td>For women with WHO Group II ovulation disorders [PCOS] who are known to be resistant to clomiphene citrate, consider one of the following second-line treatments, depending on clinical circumstances and the woman's preference:&lt;br&gt;• laparoscopic ovarian drilling or&lt;br&gt;• combined treatment with clomiphene citrate and metformin if not already offered as first-line treatment or&lt;br&gt;• gonadotropins&lt;br&gt;Women prescribed metformin should be informed of the side effects associated with its use (such as nausea, vomiting and other gastrointestinal disturbances).</td>
<td>Combination therapies involving CC and other agents (metformin, glucocorticoids, and exogenous gonadotropins) may be effective when treatment with CC alone fails to induce ovulation. There is no evidence for improved live birth rates or decreased pregnancy complications with the use of metformin either before conception or during pregnancy (Level A). Note there is a growing body of evidence on its use both for PCOS patients who get treatment and for patients with gestational diabetes.</td>
</tr>
<tr>
<td>Intervention</td>
<td>Evidence Report Findings</td>
<td>NICE</td>
<td>ASRM</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Aromatase inhibitors</td>
<td>Letrozole has a higher live birth rate than clomiphene citrate alone and lower multiple births (moderate SOE for both outcomes), with no difference in ectopic pregnancy, miscarriage, low birthweight, or time to pregnancy (low SOE for these outcomes)</td>
<td>In women with polycystic ovary syndrome, letrozole appears to be associated with a higher live birth rate, lower rates of multiple pregnancy and lower incidence of ovarian hyperstimulation syndrome (OHSS) than clomiphene citrate.</td>
<td></td>
</tr>
<tr>
<td>Surgical Management</td>
<td>There was no difference between laparoscopic ovarian drilling (LOD) and oral agents for live birth (moderate SOE) or miscarriage rates (low SOE). Multiple births were reduced given LOD (moderate SOE)</td>
<td>For women with WHO Group II ovulation disorders [PCOS] who are known to be resistant to clomiphene citrate, consider one of the following second-line treatments, depending on clinical circumstances and the woman's preference: • LOD or • combined treatment with clomiphene citrate and metformin if not already offered as first-line treatment or • gonadotropins</td>
<td>No specific recommendations</td>
</tr>
<tr>
<td>Gonadotropins alone</td>
<td>Use of gonadotropins as primary therapy does not improve outcomes compared to oral agents</td>
<td>Women with polycystic ovary syndrome who are being treated with gonadotropins should not be offered treatment with gonadotropin-releasing hormone agonist concomitantly because it does not improve pregnancy rates, and it is associated with an increased risk of ovarian hyperstimulation.</td>
<td>No specific recommendations</td>
</tr>
<tr>
<td>Adjuncts to gonadotropins</td>
<td>Pretreatment with metformin prior to ART may improve live birth rates and decrease OHSS. Use of GnRH antagonists as part of the controlled ovarian hyperstimulation protocol in IVF/ICSI reduces the incidence of OHSS compared to GnRH agonists.</td>
<td>The use of adjuvant growth hormone treatment with gonadotropin-releasing hormone agonist and/or human menopausal gonadotropin during ovulation induction in women with polycystic ovary syndrome who do not respond to clomiphene citrate is not recommended because it does not improve pregnancy rates.</td>
<td>No specific recommendations</td>
</tr>
</tbody>
</table>

*Sections or statements shaded in gray represent outcomes where our review’s findings were different from those of the major guidelines/recommendations.

Abbreviations: ART=assisted reproductive technology; ASRM=American Society for Reproductive Medicine; BMI=body mass index; CC=clomiphene citrate; GnRH=gonadotropin-releasing hormone; ICSI=intracytoplasmic sperm injection; IVF=in vitro fertilization; KQ=Key Question; NICE=National Institute for Health and Care Excellence (UK); OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome; SOE=strength of evidence; WHO=World Health Organization
Table 36. Report findings and major guidelines/recommendations—KQ 2, endometriosis*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical management</td>
<td>Strength of evidence was insufficient for all comparisons.</td>
<td>Medical treatment of minimal and mild endometriosis diagnosed as the cause of infertility in women does not enhance fertility and should not be offered.</td>
<td>There is no evidence that medical treatment of endometriosis improves fertility.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In younger women (under age 35 years) with stage I/II endometriosis-associated infertility, expectant management or superovulation with IUI can be considered as first-line therapy. For women 35 years of age or older, more aggressive therapy (superovulation with IUI or IVF) may be considered.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>IVF success rates in women with endometriosis appear to be diminished compared to women with tubal factor infertility; however, IVF likely maximizes cycle fecundity for those with endometriosis.</td>
</tr>
<tr>
<td>Surgery</td>
<td>Strength of evidence was insufficient for all comparisons.</td>
<td>Women with minimal or mild endometriosis who undergo laparoscopy should be offered surgical ablation or resection of endometriosis plus laparoscopic adhesiolysis because this improves the chance of pregnancy.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Women with ovarian endometriomas should be offered laparoscopic cystectomy because this improves the chance of pregnancy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Women with moderate or severe endometriosis should be offered surgical treatment because it improves the chance of pregnancy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Post-operative medical treatment does not improve pregnancy rates in women with moderate to severe endometriosis and is not recommended.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The benefit of laparoscopic treatment of minimal or mild endometriosis is insufficient to recommend laparoscopy solely to increase the likelihood of pregnancy.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>There is insufficient evidence to indicate that resection of endometriomas prior to IVF improves outcomes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In women with stage III/IV endometriosis-associated infertility, conservative surgical therapy with laparoscopy or possible laparotomy may be beneficial. Surgical management of an endometrioma should include resection or ablation, rather than drainage, with resection preferred. For women with stage III/IV endometriosis who fail to conceive following conservative surgery or because of advancing reproductive age, IVF-ET is an effective alternative.</td>
</tr>
</tbody>
</table>

*Sections or statements shaded in gray represent outcomes where our review’s findings were different from those of the major guidelines/recommendations.

Abbreviations: SRM=American Society for Reproductive Medicine; CC=clomiphene citrate; ET=embryo transfer; IUI=intruterine insemination; IVF=in vitro fertilization; KQ=Key Question; NICE=National Institute for Health and Care Excellence (UK)
Table 37. Report findings and major guidelines/recommendations—KQ 3, unexplained infertility*

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oral ovarian stimulation</td>
<td>There is no difference between the oral agents of letrozole and anastrozole for the outcome of ectopic pregnancy (low SOE) but evidence is insufficient for other outcomes of interest. There is no difference between differing adjunct treatments used in combination with oral agents and IUI for the outcomes of live birth, miscarriage, and OHSS (low SOE for all outcomes)</td>
<td>Do not offer oral ovarian stimulation agents (such as clomiphene citrate, anastrozole or letrozole) to women with unexplained infertility.</td>
<td>The treatment effects with non-ART treatment for unexplained infertility generally are small. Empiric treatment may do no more than hasten conception in those couples who would conceive eventually without treatment.297</td>
</tr>
<tr>
<td>Clomiphene citrate</td>
<td>There is no difference between letrozole and clomiphene for outcomes of multiple births or miscarriage (moderate SOE)</td>
<td>Inform women with unexplained infertility that clomiphene citrate as a stand-alone treatment does not increase the chances of a pregnancy or a live birth.</td>
<td>Level I evidence from randomized clinical trials supports short-term use of IUI, CC, gonadotropins and IUI, and ART treatment for unexplained infertility but is insufficient for conclusions regarding CC/IUI treatment. Clomiphene citrate treatment combined with intercourse does not increase cycle fecundity in couples with unexplained infertility compared with expectant management; however, clomiphene plus IUI does increase fecundity compared to expectant management.295</td>
</tr>
<tr>
<td>IVF</td>
<td>There are no differences between immediate IVF versus other treatments prior to IVF for the outcomes of live birth, multiple births, ectopic pregnancy, miscarriage, low birthweight, and OHSS (low SOE for all outcomes). There is however shorter time to pregnancy with immediate IVF (moderate SOE)</td>
<td>Advise women with unexplained infertility who are having regular unprotected sexual intercourse to try to conceive for a total of 2 years (this can include up to 1 year before their fertility investigations) before IVF will be considered.</td>
<td>ART therapies are considerably more costly than CC and IUI. • Adverse effects of ART and ovarian stimulation include multiple pregnancy and ovarian hyperstimulation. • When considering treatment options for couples with unexplained infertility, it is prudent to consider simple treatment before complex treatment and to balance what is known about effectiveness against the cost and adverse effects of different treatments. Offer IVF treatment to women with unexplained infertility who have not conceived after 2 years (this can include up to 1 year before their fertility investigations) of regular unprotected sexual intercourse.</td>
</tr>
</tbody>
</table>

*Sections or statements shaded in gray represent outcomes where our review’s findings were different from those of the major guidelines/recommendations.

Abbreviations: ART=assisted reproductive technology; ASRM=American Society for Reproductive Medicine; CC=clomiphene citrate; FSH=follicle-stimulating hormone; ICSI=intracytoplasmic sperm injection; IUI=intruterine insemination; IVF=in vitro fertilization; KQ=Key Question; NICE=National Institute for Health and Care Excellence (UK); OHSS=ovarian hyperstimulation syndrome; SOE=strength of evidence
### Table 38. Report findings and major guidelines/recommendations—KQ 4, tubal and peritoneal factor infertility

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM&lt;sup&gt;TM&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td>Strength of evidence was insufficient for all comparisons.</td>
<td>Women who are not known to have comorbidities (such as pelvic inflammatory disease, previous ectopic pregnancy or endometriosis) should be offered HSG to screen for tubal occlusion because this is a reliable test for ruling out tubal occlusion, and it is less invasive and makes more efficient use of resources than laparoscopy. [2004] Where appropriate expertise is available, screening for tubal occlusion using hysterosalpingo-contrast-ultrasonography should be considered because it is an effective alternative to hysterosalpingography for women who are not known to have comorbidities. [2004] Women who are thought to have comorbidities should be offered laparoscopy and dye so that tubal and other pelvic pathology can be assessed at the same time. [2004]</td>
<td>There is good evidence to support HSG as the standard first line test to assess tubal patency, but it is limited by false positive diagnoses of proximal tubal blockage. *• The evidence is fair to recommend tubal cannulation for proximal tubal obstruction in young women with no other significant infertility factors. • The evidence is fair to recommend laparoscopic fimbrioplasty or neosalpingostomy for the treatment of mild hydrosalpinges in young women with no other significant infertility factors.</td>
</tr>
<tr>
<td>Intervention</td>
<td>Evidence Report Findings</td>
<td>NICE</td>
<td>ASRM&lt;sup&gt;798&lt;/sup&gt;</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Surgery</td>
<td>Strength of evidence was insufficient for all comparisons.</td>
<td>For women with mild tubal disease, tubal surgery may be more effective than no treatment. In centres where appropriate expertise is available it may be considered as a treatment option. [2004]</td>
<td>There are no adequate trials comparing pregnancy rates with tubal surgery vs. IVF. However, IVF has a higher per-cycle pregnancy rate.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>For women with proximal tubal obstruction, selective salpingography plus tubal catheterisation, or hysteroscopic tubal cannulation, may be treatment options because these treatments improve the chance of pregnancy. [2004]</td>
<td>The evidence is fair to recommend tubal cannulation for proximal tubal obstruction in young women with no other significant infertility factors.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The evidence is fair to recommend laparoscopic fimbrioplasty or neosalpingostomy for the treatment of mild hydrosalpinges in young women with no other significant infertility factors.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tubal anastomosis for reversal of tubal sterilization has a significantly higher cumulative pregnancy rate than IVF, and it is more cost efficient, even in women 40 years of age or older.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Laparoscopic salpingectomy or proximal tubal ligation overcomes the detrimental effect of hydrosalpinges on IVF pregnancy rates in patients who are not candidates for corrective tubal surgery.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>There is good evidence for recommending laparoscopic salpingectomy or proximal tubal occlusion in cases of surgically irreparable hydrosalpinges to improve IVF pregnancy rates.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>There is good evidence to support the recommendation for microsurgical anastomosis for tubal ligation reversal.</td>
</tr>
</tbody>
</table>

Abbreviations: ART=assisted reproductive technology; ASRM=American Society for Reproductive Medicine; HSG=hysterosalpingography; ICSI=intra-cytoplasmic sperm injection; IVF=<em>in vitro</em> fertilization; KQ=Key Question; NICE=National Institute for Health and Care Excellence (UK)
Table 39. Report findings and major guidelines/recommendations—KQ 5, male factor infertility

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonadotropins</td>
<td>No relevant findings</td>
<td>Men with hypogonadotrophic hypogonadism should be offered gonadotrophin drugs because these are effective in improving fertility.</td>
<td>No specific recommendations</td>
</tr>
<tr>
<td>Surgical</td>
<td>No relevant findings</td>
<td>Where appropriate expertise is available, men with obstructive azoospermia should be offered surgical correction of epididymal blockage because it is likely to restore patency of the duct and improve fertility. Surgical correction should be considered as an alternative to surgical sperm recovery and IVF. Men should not be offered surgery for varicoceles as a form of fertility treatment because it does not improve pregnancy rates.</td>
<td>• Infertility due to obstructive azoospermia may be treated effectively by surgical reconstruction or by retrieval of sperm from the epididymis or testis, followed by IVF/ICSI. • When obstructive azoospermia results from a vasectomy performed less than 15 years before and there are no coexisting female infertility factors, microsurgical reconstruction of the reproductive tract generally is preferred over sperm retrieval and IVF/ICSI.</td>
</tr>
<tr>
<td>ART</td>
<td>Live birth rate (moderate SOE) and miscarriage (low SOE) did not differ between intracytoplasmic sperm injection (ICSI) and intracytoplasmic morphological sperm injection (IMSI).</td>
<td>Couples should be informed that ICSI improves fertilisation rates compared to IVF alone, but once fertilisation is achieved the pregnancy rate is no better than with IVF.</td>
<td>No specific recommendations</td>
</tr>
</tbody>
</table>

*Sections or statements shaded in gray represent outcomes where our review’s findings were different from those of the major guidelines/recommendations.

Abbreviations: ASRM=American Society for Reproductive Medicine; ICSI=intracytoplasmic sperm injection; IVF=in vitro fertilization; KQ=Key Question; NICE=National Institute for Health and Care Excellence (UK)
### Table 40. Report findings and major guidelines/recommendations—KQ 6, donors

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male donors</td>
<td>No evidence on short- or long-term outcomes</td>
<td>All potential semen donors should be offered counselling from someone who is independent of the treatment unit regarding the implications for themselves and their genetic children, including any potential children resulting from donated semen.</td>
<td>Psychological evaluation and counseling by a qualified mental health professional is strongly recommended for all sperm donors. The assessment should include a clinical interview and, where appropriate, psychological testing. Psychological consultation should be required for individuals in whom there appear to be factors that warrant further evaluation. In cases of directed donation, psychological evaluation and counseling are strongly recommended for the donor and his partner (if applicable) as well as for the recipient female and her partner (if applicable). The potential impact of the relationship between the donor and recipient should be explored. The psychological assessment also should address the potential psychological risks and evaluate for evidence of coercion (financial or emotional). It is important to ascertain whether the donor is well informed about the extent to which information about him might be disclosed and about any plans that may exist relating to future contact.</td>
</tr>
</tbody>
</table>

---

**Note:**

509
| Female donors | Limited evidence on short-term outcomes, no evidence on long-term outcomes | Before donation is undertaken, oocyte donors should be screened for both infectious and genetic diseases in accordance with the 'UK guidelines for the medical and laboratory screening of sperm, egg and embryo donors' (2008). Oocyte donors should be offered information regarding the potential risks of ovarian stimulation and oocyte collection. Oocyte recipients and donors should be offered counselling from someone who is independent of the treatment unit regarding the physical and psychological implications of treatment for themselves and their genetic children, including any potential children resulting from donated oocytes. | Psychological evaluation and counseling by a qualified mental health professional is strongly recommended for the oocyte donor and her partner (if applicable). The assessment should include a clinical interview and, where appropriate, psychological testing. Psychological consultation should be required for individuals in whom there appear to be factors that warrant further evaluation. In circumstances involving known donors, psychological evaluation and counseling is strongly recommended for the donor and her partner, if applicable, as well as for the recipient and her partner, if applicable. The potential impact of the relationship between the donor and recipient should be explored. The psychological assessment also should address the potential psychological risks and evaluate for evidence of coercion (financial or emotional). It is important to ascertain whether the donor is well informed about the extent to which information about her may be disclosed and about any plans that may exist relating to future contact. • All oocyte donors should be advised explicitly of the risks and adverse effects of ovarian stimulation and retrieval, with such counseling documented by informed consent in the patient's permanent medical record. Oocyte donors are exposed to the risks of controlled ovarian stimulation, oocyte retrieval, and anesthesia. The risk of OHSS is estimated to occur in 1%–2% of donation cycles and may be further reduced by the use of GnRH agonists for triggering final oocyte maturation. The risk of serious acute complications associated with these procedures is small (<0.5%). As these are independent events, the cumulative risk of multiple procedures should be similarly low. The preponderance of data does not demonstrate a significant risk of future cancers. |


While the data are limited, available evidence does not suggest that oocyte donation is associated with changes in the donor's ovarian reserve. Currently, there are no clearly documented long-term risks associated with oocyte donation and as such no definitive data upon which to base absolute recommendations. However, because of the possible health risks outlined in the preceding discussion, it is prudent to limit the number of stimulated cycles for a given oocyte donor to 6.

Table 41. Report findings and major guidelines/recommendations—all KQs: long-term outcomes of treatments

<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
</table>
| Ovulation Induction Agents    | • Infertility itself is associated with an increased risk of some cancers, and the overall risk of cancer in women exposed to oral ovulation agents, gonadotropins, and IVF/ICSI is not increased after adjustment for infertility in general or specific causes, although there are some associations with less common cancers (low SOE).  
  • For children born after ART, ICSI may be associated with an increased risk of neurodevelopmental disorders compared to IVF, although evidence is inconsistent (low SOE). | Inform women who are offered ovulation induction or ovarian stimulation that:  
  • No direct association has been found between these treatments and invasive cancer and  
  • No association has been found in the short- to medium-term between these treatments and adverse outcomes (including cancer) in children born from ovulation induction and  
  • Information about long-term health outcomes in women and children is still awaited. | No specific recommendations |
<table>
<thead>
<tr>
<th>Intervention</th>
<th>Evidence Report Findings</th>
<th>NICE</th>
<th>ASRM</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVF/ICSI</td>
<td>• Infertility itself is associated with an increased risk of some cancers, and the overall risk of cancer in women exposed to oral ovulation agents, gonadotropins, and IVF/ICSI is not increased after adjustment for infertility in general or specific causes, although there are some associations with less common cancers (low SOE). In one large study from the United Kingdom, the risk of colorectal cancer was not increased among women undergoing ART compared to the general population, but was increased compared infertile women who did not undergo ART (low SOE). • For children born after ART, ICSI may be associated with an increased risk of neurodevelopmental disorders compared to IVF, although evidence is inconsistent (low SOE).</td>
<td>Inform women that while the absolute risks of long-term adverse outcomes of IVF treatment, with or without ICSI, are low, a small increased risk of borderline ovarian tumours cannot be excluded. Inform people who are considering IVF treatment that the absolute risks of long-term adverse outcomes in children born as result of IVF are low.</td>
<td>No specific recommendations</td>
</tr>
</tbody>
</table>

Abbreviations: ART=assisted reproductive technology; ASRM=American Society for Reproductive Medicine; ICSI=intracytoplasmic sperm injection; IVF= in vitro fertilization; KQs=Key Questions; NICE=National Institute for Health and Care Excellence (UK)
In general, this review’s findings are consistent with the guidelines cited above—there is a general consensus that the overall body of evidence for many aspects of infertility treatment across all patient groups is limited. One consistent limitation is the relative paucity of studies utilizing live birth per couple as the primary outcome. Where there are differences between our findings and these guidelines, or between guidelines, they can be attributed to the following:

- Differences in study inclusion/exclusion criteria, particularly in study dates and outcomes considered.
- Differences in grading individual study quality, which are then reflected in grading of the overall strength of evidence.
- Differences in the approach to guideline development—the NICE guidance follows the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system similar to our approach, while the ASRM uses a less structured approach.
- Differences in settings, which affect the manner in which cost considerations are weighed in formulating recommendations. NICE guidance is designed for use within the UK National Health Service, and cost-effectiveness is considered from a health system perspective. Conversely, the ASRM guidelines are meant to be applied in the more diverse U.S. setting, where coverage of infertility services is much more varied, and the perspective of individual payers and patients, who are likely to bear a much higher proportion of the cost of care, must be considered.

### Applicability

Table 42 summarizes the applicability scores across KQs.

<table>
<thead>
<tr>
<th>Issues</th>
<th>KQ 1 N=56</th>
<th>KQ 2 N=7</th>
<th>KQ 3 N=50</th>
<th>KQ 4 N=8</th>
<th>KQ 5 N=23</th>
<th>KQ 6 N=5</th>
<th>Across All KQs N=21</th>
<th>Total N=151</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Study population demographics not representative of intended population</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Narrow or unrepresentative severity/stage/comorbidity</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Intervention</td>
<td>Treatment protocol not representative of current practice</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Change in standard of care</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Comparator</td>
<td>Comparator not representative of current practice</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Timing of outcome assessment</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Setting</td>
<td>Standards or access to care vary from U.S. setting</td>
<td>32</td>
<td>5</td>
<td>27</td>
<td>3</td>
<td>11</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>
Issues | KQ 1 N=56 | KQ 2 N=7 | KQ 3 N=50 | KQ 4 N=8 | KQ 5 N=23 | KQ 6 N=5 | Across All KQs N=21 | Total N=151
---|---|---|---|---|---|---|---|---
Specialty population or level of care | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0

Numbers in cells represent the number of included studies that were identified as having potential issues related to the specific item. Columns represent numbers for each Key Question and then for all included studies.

Two broad issues relate to the overall applicability of the available evidence to clinical practice in the United States—one geographic and one temporal. Many of the randomized controlled trials (RCTs) meeting our criteria were performed outside of the United States. Leaving aside any issues related to differences in study oversight or reporting, the populations of these studies may differ from U.S. infertility patients in two potentially important ways.

The first issue is that there may be clinically relevant differences between populations in terms of non-clinical factors affecting outcomes. For example, live birth rates for African-American women undergoing ART in the US are lower than for white women, which may reflect issues related to socioeconomic status, insurance coverage, or other factors (such as well-established racial differences in the risk of many adverse pregnancy outcomes). Differences in access to infertility services between countries may lead to differences in the likelihood of treatment success. Although the estimate of any relative difference between two interventions derived from an unbiased RCT should in theory be independent of the probability of specific outcomes, the more clinically relevant absolute difference may be substantially different (e.g., the risk of preterm birth in African-American compared to white women is consistently elevated). To the extent that the probability of specific outcomes of interest may differ between populations because of differences in genetic risk, exposures to other factors affecting risk, or non-biological factors such as access to care, there may be substantial differences in estimates of absolute risk differences. For relatively uncommon but important outcomes, these differences might also affect precision of estimates—confidence intervals for any treatment effect will be wider in populations where the outcome is less common.

In addition to the potential impact of race/ethnicity, there may be important differences in the distribution of socioeconomic status between populations. Access to infertility diagnosis and treatment varies across countries, and certainly within the United States. Differences in socioeconomic status could affect applicability in several ways. Differences in access to care may lead to differences in the spectrum of severity of “disease” for U.S. patients who given the financial burden of treatment options they may wait longer to undergo evaluations. Although summary statistics of baseline characteristics may allow some judgment of comparability, there may be potentially important differences in the distribution that are obscured by the typical reporting of means and standard deviations (particularly if the underlying characteristic is not normally distributed), or by differences within a given stage. Socioeconomic status may also potentially affect some important outcomes independently of any specific treatment—for example, neurodevelopmental outcomes such as specific learning skills may be strongly correlated with parental socioeconomic status.

The second issue is that changes in practice over time have a major impact on applicability, particularly for long-term outcomes. The long lag time between exposure to infertility treatment and the potential development of longer term outcomes such as cancer means that data available today necessarily reflect women exposed to treatments at least 10 years in the past; even if the specific exposure is similar, there may be differences between past and current practice in potentially important attributes such as dosage, timing, patient selection criteria, use of
adjunctive treatments, etc. For example, evidence that immediate use of IVF leads to shorter time to pregnancy than strategies where IVF is used only after a trial of agents such as clomiphene or gonadotropins has led to a change in guidelines, which now suggest that the cumulative exposure to gonadotropins during the course of treatment is likely to decrease compared to earlier cohorts of women, reducing any long-term risks.

In addition, there may be cohort effects in terms of other exposures that may affect the absolute risk of some outcomes (e.g., changes in the use of postmenopausal hormone replacement therapy or ages of mammography screening affecting breast cancer risk), which in turn would impact any additional absolute risk due to exposure to infertility treatments. Because of this phenomenon, there is likely to always be some unresolvable uncertainty about long-term outcomes for both parents undergoing current infertility treatments and their children.

Implications for Clinical and Policy Decision Making

Our review found considerable uncertainty about the comparative effectiveness of many available treatments for infertility, both those targeted at specific causes and those like ART that are used broadly in patients with a wide range of underlying diagnoses. A large component of this uncertainty is due to inconsistency in the choice of reported outcomes. This is especially true of the choice for primary measure of “success.” Although there is a growing consensus that live birth rate per couple (either per cycle or per a number of treatment cycles) undergoing a particular intervention is the most relevant outcome for clinical decision making, many studies, particularly older ones, report intermediate outcomes such as clinical or ongoing pregnancy rates. There is some evidence that conclusions about relative effectiveness are similar whether ongoing pregnancy or live birth is used as the primary outcome. However, there may be relevant differences in absolute effectiveness, as discussed under Limitations. Although inclusion of studies using these intermediate outcomes would have certainly increased the potential pool of studies, our Technical Expert Panel felt that limiting included studies to those that reported live birth was important clinically. From a study design and feasibility perspective, intermediate outcomes reduce costs through smaller sample size requirements and shorter overall time to reaching endpoints (discussed in more detail under Research Recommendations), but feasible studies that do not sufficiently resolve uncertainty about specific clinical or policy decisions to allow confidence in those decisions are ultimately not an efficient use of resources.

Another source of uncertainty is a lack of consensus or clarity on the relative importance of different outcomes to patients, clinicians, and policymakers. Many clinical decisions regarding infertility treatment involve tradeoffs between, at least, the number of treatment cycles required, the cumulative probability of a successful outcome, and the relative probability of multiple gestations. Depending on the health system, these clinical tradeoffs can have financial implications for patients as well. There is a striking lack of evidence on the relative value patients place on different outcomes related to infertility treatment, with relatively few studies using standard methods of preference elicitation available. This makes decision analyses, both clinical and economic, difficult. This lack of data also inhibits design and testing of policies intended to optimize outcomes (e.g., insurance coverage for elective single-embryo transfer over multiple cycles for eligible patients in order to take away incentives for multiple-embryo transfer when patients pay out of pocket).

As discussed above in the applicability section, the combination of continuous changes in infertility treatment and the long time horizon needed to obtain evidence on long-term safety outcomes for both parents and children means that there will always be a degree of unresolvable
uncertainty about long-term safety for patients making decisions now. To some extent, this is true of most clinical decision making (for example, estimates of the impact of specific interventions such as cancer screening on life expectancy are based on assumptions about treatment effectiveness, competing risks, etc., that do not reflect potential future changes). This is another area where more insight into relative preferences for both outcomes and timing of outcomes (e.g., is there a potential increased risk in the risk of cancer in 20 or 30 years that patients would be willing to trade off for an increased probability of a live birth within the next year) would be helpful.

A potential first step toward addressing these uncertainties within the U.S. context would be to create a formal structure for evaluating specific decisions relevant to infertility management, similar to the GRADE approach used by NICE. Achieving consensus on at least the relative importance of specific outcomes (potentially including costs), the ideal method for measuring those outcomes, and acceptable trade-offs between harms and benefits (such as the number of cases of OHSS per live birth, or number of preterm births attributable to multiple births per live birth), would be useful for structuring future reviews, guidelines development, and, through methods such as value-of-information analysis, prioritizing future research.

**Limitations of the Systematic Review Process**

Several aspects of the review process may have affected the results. First, there were constraints in our search strategy, developed in consultation with the Key Informants and TEP. We did not review evidence on the diagnostic evaluation of patients with infertility. Although this is obviously a critical question for patients, clinicians, and policymakers, it was outside of the scope of the review.

We limited the search to papers published after the cutoff date of the previous AHRQ Evidence Report on ART. This meant that studies completed prior to the cutoff date, which otherwise might have met inclusion criteria, were excluded. While we believe that the majority of these studies were included in the systematic reviews we used to supplement each KQ, it is possible that potentially relevant articles were missed. Given broad changes in clinical practice (such as increasing use of ICSI even in the absence of specific male factor infertility and decreasing use of empiric gonadotropin therapy prior to ART in couples with unexplained infertility) over the past decade, the impact of missing earlier studies on conclusions about comparative effectiveness of currently used treatment alternatives is unclear.

We limited the outcomes to those considered most important by key stakeholders, using a formal prioritization process described in the Methods section, in an attempt to keep the scope of the review tractable. We specifically limited the review to articles that reported live birth as the primary pregnancy-related outcome, excluding studies that reported pregnancy rates alone (including studies reporting clinical or ongoing pregnancies). There is growing consensus that live birth is the most appropriate outcome for studies of infertility treatment effectiveness, particularly when expressed as the cumulative probability of live birth per couple over time rather than on a per-cycle basis, since this is the most clinically relevant information for a given couple. In 2010, the Cochrane group found that live birth is still infrequently reported in trials of infertility treatment, but there is some evidence suggesting that, in studies that report both clinical pregnancy and live birth outcomes, the magnitude and direction of effect are similar. Including otherwise eligible studies that reported clinical pregnancy rates alone may have provided some additional evidence relevant to the KQs. However, because the use of a surrogate or intermediate outcome such as clinical pregnancy rate affects the strength of
Our review process was structured by KQ and, for each outcome, we required that outcomes be reported for the specific patient population covered by the KQ, or, in the case of observational studies, that the underlying diagnosis be included in multivariate analyses so that the reported overall measure of association accounted for variability based on diagnosis. This approach led to the exclusion of a number of studies of treatments used across multiple diagnoses, particularly those involving ART. The extent to which specific outcomes might differ based on underlying diagnosis is unclear. For long-term outcomes such as cancer in female patients, there is evidence that certain infertility diagnoses increase risk independently of any treatment effects (for example, PCOS and endometrial cancer, or endometriosis and ovarian cancer). Risks of some adverse outcomes (e.g., OHSS or ectopic pregnancy) differ in different populations (e.g., women with PCOS, tubal factor, or endometriosis infertility). However, there is less evidence that treatment effectiveness varies by diagnosis, although even when relative differences are similar, there still may be clinically important differences in the absolute probability of specific benefits and harms. One alternative approach to structuring the review would be to focus diagnosis-specific reviews only on treatments used prior to initiation of ART, and report comparative effectiveness of treatments used in ART under the assumption that outcomes are similar across patient populations. However, even if this approach expanded the evidence base, there still would be residual uncertainty surrounding quantitative estimates of outcome likelihood in specific patient populations.

Last, we did not include studies published in languages other than English, primarily due to resource limitations. However, given differences in the way infertility evaluation and treatment is financed in different countries, our judgment (discussed in more detail under Applicability) is that there may be important differences, both measurable and unmeasurable, between couples undergoing infertility in the United States compared with other countries. Inclusion of non-English language studies would only make this problem worse.

Research Recommendations

In an era of constrained resources, future clinical research, especially comparative effectiveness research—which helps resolve current uncertainties regarding clinical or policy decisions—should receive priority. For most of the KQs, there are multiple areas of remaining uncertainty based on the existing evidence. In part because of the diversity of causes and treatment options, it is difficult to make specific recommendations for specific topics.

Before setting a specific agenda for future research in infertility, we believe a more general approach to identifying priorities would be helpful. Achieving consensus on the relative priority of specific outcomes, incorporating the perspective of multiple stakeholders (similar to the approach used for developing a research agenda for comparative effectiveness research for uterine fibroids.311,312 Ideally, these outcome priorities would be used for subsequent evidence syntheses and guideline development.

As part of this consensus process, additional areas of discussion include:

- Formal consideration of the limits of acceptability for specific quantitative harms (e.g., preterm birth) and clinically meaningful differences in benefits (e.g., live birth).
- Formal discussion of the potential role of cost-effectiveness in decision making, including issues of willingness-to-pay and appropriate choice of outcome. This is particularly important because there are significant methodological challenges to the use...
of “standard” measures such as quality-adjusted life expectancy in the setting of infertility treatment.

- Issues related to study design, particularly from the patient stakeholder perspective. For example, in settings where patients and/or clinicians may have strong preferences for specific treatments, recruitment into RCTs may be difficult.\(^{312}\) In the uterine fibroid consensus process, patient stakeholders strongly preferred observational designs to randomized treatment assignment.\(^{311}\) Discussion of potential trade-offs between risk of bias, efficiency, ability to measure all relevant potential confounders and effect modifiers, appropriateness of alternative approaches such as Zelen randomization (where subjects are randomized prior to consent, then allowed to either receive the assigned treatment or choose the alternative\(^{314}\)), and the likelihood that a specific study design would resolve a specific area of uncertainty should all be included.

- Issues related to data reporting. Particularly for ART and other treatments which are used for multiple indications, reporting of results separately by indication in both randomized trials and large observational studies would be extremely useful. Although these subgroup results may have insufficient power to detect clinically relevant differences within the context of individual studies (particularly RCTs), their routine publication would eventually allow synthesis of results using methods such as meta-analysis (including individual-level meta-analysis.)

Part of this process could include value-of-information analysis, a formal method for quantifying the impact of existing uncertainty on the likelihood of making the “wrong” decision.\(^{315-317}\) Although the approach has classically been used in the framework of cost-effectiveness, the basic methods for illustrating the impact of uncertainty on the probability of making an optimal decision can also be applied using specific harms and benefits.

In addition to development of a specific consensus-driven approach to resolving uncertainty, other specific recommendations apply across all areas of infertility treatment. Empiric measurement of patient preferences using validated measures would have substantial impact. In the context of infertility treatment, where approaches such as the standard gamble or time trade-off (which require trading off risk of immediate death or life expectancy versus specific health benefits) may be both conceptually difficult and counter-intuitive to patients, approaches such as discrete choice experiments (DCE) may be preferable (DCE, or conjoint analysis, also has the advantage of being able to explicitly incorporate costs to measure willingness-to-pay).\(^{318,319}\)

The SART CORS and National ART Surveillance System (which includes data submitted through SART CORS (the majority of clinics providing ART) as well as a smaller number of non-SART participating clinics who report directly to CDC) databases are outstanding examples of what a large-scale, population-based registry can achieve in terms of providing data on treatment outcomes. However, the major limitation of the database in the past has been that data are only published on a per-cycle, rather than per-couple, basis. Recently the database methods have changed and now they are publicly reporting the cumulative success rate per patient. Results, however, are still reported at the clinic level, so patients who receive care at more than one clinic do not have the full range of outcomes captured, and there is no mechanism for prospectively collecting long-term outcomes of patients or children. Facilitating reporting of results so that outcomes are reported on a per-couple basis will substantially improve the ability to generate estimates of the likely outcome of specific ART-related decisions.
Based on input from key informants and our Technical Expert Panel, we structured the review based on infertility diagnosis, and required studies to report outcomes specifically by diagnosis, or to adjust for diagnosis in multivariable analyses. As noted above, this led to exclusion of a number of papers, particularly those related to ART methods. There is clear evidence that the probability of some outcomes of interest, both short-term (e.g., OHSS) and long-term (certain cancers) differs based on underlying diagnosis. Although this may not be the case for all outcomes, we believe it would be helpful for future studies of interventions performed in patients with different underlying diagnoses to report results separately by diagnosis. Within an individual study powered on the basis of the total patients, estimates of diagnosis-specific outcomes may be too imprecise to confidently rule out clinically relevant differences—consistency of reporting would allow formal synthesis of estimates across studies.

We found very limited evidence on outcomes among sperm or oocyte donors. Oocyte donors, who undergo controlled ovarian hyperstimulation and oocyte retrieval in the same manner as patients undergoing IVF using their own eggs, have, in theory, at least the same risk of short-term adverse events as patients. The frequency with which oocyte donors are used is increasing, and evidence from the SART CORS database suggests that the risk of certain pregnancy complications is lower when donor oocytes are used.37,38 If demand for donor oocytes continues to increase, much more evidence on the specific short- and long-term outcomes of donation (especially if a donor undergoes multiple cycles) is needed.

**Conclusion**

Recently there has been growing adaptation of more rigorous methods for evaluating treatments for infertility, particularly regarding treatments for PCOS and approaches to timing of interventions in patients undergoing ART. In addition, ongoing refinements to the SART CORS database continue to make it a valuable resource, particularly for data on short-term outcomes. However, given the diversity of infertility causes and treatments, there is considerable residual uncertainty about the optimal treatment options for specific patients. Consensus on which outcomes to report (such as encouraging reporting of live birth rates on a per couple basis as well as per cycle, and, for studies of treatment such as ART, reporting of both overall and diagnosis-specific outcomes) and which areas of uncertainty are most important to resolve (in order to prioritize research) is needed to improve the ability of patients and clinicians to make decisions about the most appropriate treatment.
References


Acronyms and Abbreviations

AHRQ  Agency for Healthcare Research and Quality
ART  assisted reproductive technology
ASRM  American Society for Reproductive Medicine
BMI  body mass index
CDC  Centers for Disease Control and Prevention
CI  confidence interval
DCE  discrete choice experiments
DHEA  dehydroepiandrosterone
DOR  diminished ovarian reserve
EHC  Effective Health Care
EPC  Evidence-based Practice Center
eSET  elective single-embryo transfer
FASTT  Fast Track and Standard Treatment
FORT-T  Forty and Over Treatment Trial
FSH  follicle-stimulating hormone
GnRH  gonadotropin-releasing hormone
GRADE  Grading of Recommendations Assessment, Development, and Evaluation
hCG  human chorionic gonadotropin
hMG  human menopausal gonadotropin
HR  hazard ratio
ICTRP  International Clinical Trials Registry Platform
ICSI  intra-cytoplasmic sperm injection
IMSI  intra-cytoplasmic morphologically selected sperm injection
IUI  intrauterine insemination
IVF  in vitro fertilization
IVM  in vitro maturation
KIs  Key Informants
KQ  Key Question
NASS  National ART Surveillance System
NICE  National Institute for Health and Care Excellence (UK)
NIH  National Institutes of Health
NR  not reported
NS  not statistically significant
OAT  oligo-astheno-teratozoospermia
OHSS  ovarian hyperstimulation syndrome
OR  odds ratio
PCOS  polycystic ovary syndrome
PGS  preimplantation genetic screening
PICOTS  populations, interventions, comparators, outcomes, timing, settings
PICS  physiological intra-cytoplasmic sperm injection
PRISMA  Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT  randomized controlled trial
rFSH  recombinant follicle-stimulating hormone
RR  relative risk
SART  Society for Assisted Reproductive Technology
SART CORS  Society for Assisted Reproductive Technology Clinic Outcome Reporting System
SD  standard deviation
SIR  standardized incidence ratio
SOE  strength of evidence
SRC  Scientific Resource Center
TEP  Technical Expert Panel
TESE  extracted testicular sperm
uFSH  urinary follicle-stimulating hormone
UK  United Kingdom
WHO  World Health Organization
# Appendix A. Exact Search Strings

### PubMed® Search Strategy (October 3, 2018)

<table>
<thead>
<tr>
<th>Set</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
<td>Terms</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>#4</td>
<td>#1 AND #2 AND #3</td>
</tr>
<tr>
<td>#5</td>
<td>Dates: 2007/01/01 – present</td>
</tr>
<tr>
<td>#6</td>
<td>Limit: English</td>
</tr>
</tbody>
</table>

**Embase® Search Strategy (October 3, 2018)**

Platform: Embase.com  

<table>
<thead>
<tr>
<th>Set</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>’infertility’[exp OR ’anovulation’[exp OR ’infertility”:ab,ti OR ’infertile”:ab,ti OR ’subfertility”:ab,ti OR ”subfertility”:ab,ti OR ”sub-fertility”:ab,ti OR ”sub-fertile”:ab,ti OR ”anovulation”:ab,ti OR ”aspermia”:ab,ti OR ”asthenozoospermia”:ab,ti OR ”azoospermia”:ab,ti OR ”oligospermia”:ab,ti OR ”sertoli cell-only syndrome”:ab,ti</td>
</tr>
<tr>
<td>Set</td>
<td>Terms</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>#2</td>
<td>'infertility therapy'/exp OR 'diet therapy'/exp OR 'weight reduction'/exp OR 'exercise'/exp OR 'kinesitherapy'/exp OR 'fertility promoting agent'/exp OR 'clomifene'/exp OR 'gonadotropin'/exp OR 'metformin'/exp OR 'hormone antagonist'/exp OR 'gonadotropin'/exp OR 'watchful waiting'/exp OR 'family planning'/exp OR 'ovulation detection'/exp OR 'gynecologic surgery'/exp OR 'arginine'/exp OR 'aspartic acid'/exp OR 'citric acid'/exp OR 'flavonoid'/exp OR 'corticosteroid'/exp OR 'low level laser therapy'/exp OR 'dexamethasone'/exp OR 'vasosvasostomy'/exp OR 'urofollitropin'/exp OR 'electrocoagulation'/exp OR 'prenatal diagnosis'/exp OR 'artificial insemination'/exp OR 'ovulation prediction'/exp OR 'genetic screening'/exp OR 'tetrozole'/exp OR 'cetrorelix'/exp OR 'ganirelix'/exp OR 'recombinant follitropin'/exp OR 'progesterone'/exp OR 'recombinant chorionic gonadotropin'/exp OR 'reproductive techniques':ab,ti OR 'reproductive technology':ab,ti OR 'reproductive technologies':ab,ti OR 'assisted reproductive':ab,ti OR 'IUI':ab,ti OR 'in vitro':ab,ti OR 'ivf':ab,ti OR 'natural family planning':ab,ti OR 'ovulation detection':ab,ti OR 'intrauterine insemination':ab,ti OR 'intrauterine implantation':ab,ti OR 'embryo transfer':ab,ti OR 'artificial insemination':ab,ti OR 'assisted pregnancy':ab,ti OR 'assisted reproduction':ab,ti OR 'ovulation induction':ab,ti OR 'ovarian stimulation':ab,ti OR 'ovarian hyperstimulation':ab,ti OR 'clomiphene':ab,ti OR 'serophene':ab,ti OR 'clomiphene citrate':ab,ti OR 'letrozole':ab,ti OR 'metformin':ab,ti OR 'gonadotropins':ab,ti OR 'gonadotropin releasing hormone':ab,ti OR 'hormone antagonists':ab,ti OR 'menotropins':ab,ti OR 'menopur':ab,ti OR 'repronex':ab,ti OR 'goserelin':ab,ti OR 'Zoladex':ab,ti OR 'leuprolide':ab,ti OR 'Lupron':ab,ti OR 'nafarelin':ab,ti OR 'Synarel':ab,ti OR 'cetrorelix':ab,ti OR 'Degarelix':ab,ti OR 'Firmagon':ab,ti OR 'ganirelix':ab,ti OR 'antagon':ab,ti OR 'exercise':ab,ti OR 'diet':ab,ti OR 'weight loss':ab,ti OR 'natural family planning':ab,ti OR 'timed intercourse':ab,ti OR 'Billings':ab,ti OR 'Creighton':ab,ti OR 'rhythm method':ab,ti OR 'standard days method':ab,ti OR 'calendar method':ab,ti OR 'basal body temperature method':ab,ti OR 'hysteroscopy':ab,ti OR 'hysteroscopic':ab,ti OR 'microhysterectomy':ab,ti OR 'microhysteroscopic':ab,ti OR 'ovarian drilling':ab,ti OR 'donor oocytes':ab,ti OR 'oocyte retrieval':ab,ti OR 'sperm donation':ab,ti OR 'sperm donor':ab,ti OR 'semen donation':ab,ti OR 'semen donor':ab,ti OR 'sperm extraction':ab,ti OR 'sperm retrieval':ab,ti OR 'sperm aspiration':ab,ti OR 'tesa':ab,ti OR 'micro tesi':ab,ti OR 'mesa':ab,ti OR 'pesa':ab,ti OR 'ejaculatory duct resection':ab,ti OR 'recombinant human follicle stimulating hormone':ab,ti OR 'rFSH':ab,ti OR 'rhFSH':ab,ti OR 'hormone therapy':ab,ti OR 'laser vaporization':ab,ti OR 'laser vaporisation':ab,ti OR 'dexamethasone':ab,ti OR 'vasectomy reversal':ab,ti OR 'sterilization reversal':ab,ti OR 'superovulation':ab,ti OR 'follistim':ab,ti OR 'Gonal F':ab,ti OR 'Gonal-F':ab,ti OR 'Bravelle':ab,ti OR 'crinone':ab,ti OR 'endometrim':ab,ti OR 'prometrium':ab,ti OR 'fulguration':ab,ti OR 'endometriosis excision':ab,ti OR 'endometrioma excision':ab,ti OR 'ovarian cystectomy':ab,ti OR 'tubal ligation reversal':ab,ti OR 'tubal cannulation':ab,ti OR 'therapeutic donor insemination':ab,ti OR 'ovulation prediction':ab,ti OR 'ovidrel':ab,ti OR 'assisted hatching':ab,ti OR 'preimplantation diagnosis':ab,ti OR 'preimplantation genetic diagnosis':ab,ti OR 'preimplantation screening':ab,ti OR 'preimplantation genetic screening':ab,ti OR 'preimplantation testing':ab,ti OR 'preimplantation genetic testing':ab,ti</td>
</tr>
<tr>
<td>#3</td>
<td>'randomized controlled trial'/exp OR 'crossover procedure'/exp OR 'double blind procedure'/exp OR 'single blind procedure'/exp OR 'random':ab,ti OR 'factorial':ab,ti OR 'crossover':ab,ti OR 'cross NEAR/1 over':ab,ti OR 'placebo':ab,ti OR 'doubt':ab,ti OR 'NEAR/1 blind':ab,ti OR 'sing':ab,ti OR 'NEAR/1 blind':ab,ti OR 'assign':ab,ti OR 'allocation':ab,ti OR 'volunteer':ab,ti OR 'clinical study'/exp OR 'clinical trial':ab,ti OR 'clinical trials':ab,ti OR 'evaluation':exp OR 'evaluation study':ab,ti OR 'evaluation studies':ab,ti OR 'intervention study':ab,ti OR 'intervention studies':ab,ti OR 'case control':ab,ti OR 'cohort study':exp OR 'cohort':ab,ti OR 'prospective':ab,ti OR 'prospectively':ab,ti OR 'retrospective':ab,ti OR 'follow up':exp OR 'follow up':ab,ti OR 'comparative effectiveness':exp OR 'comparative study':exp OR 'comparative study':ab,ti OR 'comparative studies':ab,ti OR 'evidence based medicine':exp OR 'systematic review':ab,ti OR 'meta-analysis':ab,ti OR 'meta-analyses':ab,ti OR 'case report':exp OR 'case study':exp OR 'editorial':exp OR 'letter':exp OR 'note':exp OR 'AND' [humans]/lim AND [embase]/lim NOT [medline]/lim</td>
</tr>
<tr>
<td>#4</td>
<td>#1 AND #2 AND #3</td>
</tr>
<tr>
<td>#5</td>
<td>#4 AND [2007-2015]/py</td>
</tr>
<tr>
<td>#6</td>
<td>#5 AND [english]/lim</td>
</tr>
</tbody>
</table>
Cochrane Search Strategy (October 3, 2018)
Platform: Wiley
Database searched: Cochrane Database of Systematic Reviews

<table>
<thead>
<tr>
<th>Set</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>MeSH descriptor Infertility explode all trees OR MeSH descriptor Anovulation explode all trees OR &quot;infertility&quot;:ab,ti,kw OR &quot;infertile&quot;:ab,ti,kw OR &quot;subfertility&quot;:ab,ti,kw OR &quot;sub-fertility&quot;:ab,ti,kw OR &quot;sub-fertile&quot;:ab,ti,kw OR &quot;aspermia&quot;:ab,ti,kw OR &quot;asthenozoospermia&quot;:ab,ti,kw OR &quot;azoospermia&quot;:ab,ti,kw OR &quot;oligospermia&quot;:ab,ti,kw OR &quot;sertoli cell-only syndrome&quot;:ab,ti,kw</td>
</tr>
<tr>
<td>#2</td>
<td>Dates: 2007/01/01 – present</td>
</tr>
<tr>
<td>#3</td>
<td>Limit: Cochrane Reviews</td>
</tr>
</tbody>
</table>

Grey Literature Searches

ClinicalTrials.gov (December 16, 2015)

<table>
<thead>
<tr>
<th>Set</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>infertility OR infertile OR subfertility OR subfertile OR sub-fertility OR sub-fertile</td>
</tr>
<tr>
<td>Limits</td>
<td>interventional studies</td>
</tr>
</tbody>
</table>

Total number of results exported: 858

Results were imported into Microsoft Excel and refined as follows:
1. Limited to studies with Completed status – 482 records removed, 376 remaining
2. Limited to studies registered from 2005 forward – 22 records removed, 354 remaining

Total number of results for screening: 354
**ClinicalTrials.gov – Narrow search for the Appendix H. Supplemental Project to Assess the Transparency of Reporting for Trials Evaluating Treatment for Infertility (February 5, 2016)**

<table>
<thead>
<tr>
<th>Set</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search terms</td>
<td>infertility OR infertile OR subfertility OR subfertile OR sub-fertility OR sub-fertile</td>
</tr>
</tbody>
</table>
| Condition terms      | polycystic ovary OR polycystic ovaries OR PCOS OR PCO endometriosis OR endometrioma  
                       | unexplained OR ovarian reserve OR DOR OR ovarian response OR POR or responded OR maternal age OR AMA OR reproductive age  
                       | tubal factor OR peritoneal factor OR pelvic adhesions OR pelvic adhesive OR hydrosalpinx  
                       | OR tubal obstruction OR tubal blockage  
                       | male factor OR male infertility OR Oligozoospermia OR Oligospermia OR Azoospermia OR Asthenospermia OR Teratospermia  
                       | oocyte donor OR oocyte donation OR egg donation OR egg donor OR sperm donor OR  
                       | sperm donation OR donor eggs OR donor oocytes OR donor sperm OR oocyte recipient |
| Limits               | interventional studies                                              |

Total number of results: 494

**WHO: International Clinical Trials Registry Platform Search Portal (January 27, 2016)**

<table>
<thead>
<tr>
<th>KQs 1-6</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>infertility OR infertile OR subfertility OR subfertile OR sub-fertility OR sub-fertile</td>
</tr>
<tr>
<td>Recruiting status</td>
<td>All</td>
</tr>
</tbody>
</table>

Total number of results exported: 1708

Results were imported into EndNote® and refined as follows:

1. Removal of records originating from ClinicalTrials.gov (the ClinicalTrials.gov database was searched separately) -- 1013 records removed, 695 remaining
2. Keyword searches to identify records containing any of the following terms of interest: birth, delivery, ectopic, miscarriage, death, cancer, OHSS, ovarian hyperstimulation syndrome, time to pregnancy, costs-- 442 records removed, 253 remaining

Total number of results for screening: 253
Appendix B. Data Abstraction Elements

Study Characteristics

- Study Identifiers
  - Study Name or Acronym
  - NCT number or other trial registry identifier
  - Last name of first author
- Additional Articles Used in This Abstraction
- Study Sites
  - Single center, Multicenter, Unclear/Not reported
  - Number of sites
- Geographic Location (Select all that apply)
  - US, Canada, UK/Europe, Latin America, Middle East (including Israel), Asia, Africa, Australia/NZ, Unclear/Not reported
- Study Design
  - RCT
  - Observational
- Funding Source (Select all that apply)
  - Government, Industry, Non-government/non-industry, Unclear/Not reported
- Setting (Select all that apply)
  - Subspecialty practice (infertility specialist, urologist, etc.); General gynecology practice; Family practice/general internist/nurse practitioner/other non-gynecologist primary care provider; Unclear/Not reported
- Study Definition of Infertility
  - No pregnancy after 12 months of regular intercourse for women <35 years old or 6 months for women 35 and older; Other (specify); Not applicable; Not reported
- Study Enrollment/Study Completion
  - N enrolled/included
  - N completed
- Key Question Applicability (Select all that apply)
  - KQ1, KQ2, KQ3, KQ4, KQ5, KQ6
- Baseline Characteristics – Record the following elements for Total Population, Women, Men, Arm 1, Arm 2, Arm 3, and Arm 4 (as applicable)
  - Number of Patients (N and %)
  - Age in years
    - Mean
    - Median
    - Standard Deviation
    - Min
    - Max
    - 25% IQR
    - 75% IQR
    - Categorical
    - Other, specify
  - Race/Ethnicity (N and %)
• Hispanic or Latino
• Black/African American
• American Indian or Alaska Native
• Asian
• Native Hawaiian or Pacific Islander
• White
• Multiracial
• Other (specify)

• Were there significant differences noted between groups in any baseline characteristic? (Yes/No)
  o If yes, please explain the differences

• Comments

**Intervention Characteristics**

• Is the comparison within a single intervention class or between classes?

• Intervention Descriptors
  o Describe the intervention received by each patient group.

• Indicate components of the intervention (For each Arm)
  o Oral Ovulation Induction with IUI
  o Oral Ovulation Induction without IUI
  o Surgical Management
  o Gonadotropins with IUI
  o Gonadotropins without IUI
  o IVF
  o ICSI
  o No intervention / expectant management

• Indicate all intervention characteristics that are varied in this study
  o IUI Details
    ▪ IUI methods
    ▪ Adjuvant treatments
  o Oral Ovulation Induction Details
    ▪ Medication type
    ▪ Timing of medication
    ▪ Adjuvant treatments
    ▪ Dose
  o Surgical Management Details
    ▪ Female – Surgical approach (e.g., laparoscopic vs. open)
    ▪ Female – Surgery vs. alternatives
    ▪ Male – Surgical repair
  o Gonadotropin Details
    ▪ Ovarian stimulation (non-IVF) – medication type
    ▪ Ovarian stimulation (non-IVF) – timing
  o IVF Details
    ▪ Pre-stimulation/adjuvant methods
    ▪ Down regulation methods
    ▪ Ovarian stimulation – medication type
Ovarian stimulation – monitoring
Ovarian stimulation – poor responders
Ovarian stimulation – natural cycle IVF
Ovulation triggering methods
Oocyte retrieval methods
Sperm retrieval methods
Laboratory phase methods
Embryo transfer – stage of development
Embryo transfer – # of embryos
Embryo transfer – transfer technique
Luteal phase support
Frozen embryos
Prevention of ovarian hyperstimulation syndrome
  o ICSI Details
    ■ Sperm retrieval methods
    ■ Sperm injection methods
    ■ ICSI vs IVF
    ■ Other (specify)
  • Comments

Outcomes
  • Select the outcome reported on this form:
    o Live Birth
      ■ Singleton (reported per cycle)
      ■ Singleton (reported per patient)
      ■ Multiple (reported per cycle)
      ■ Multiple (reported per patient)
      ■ Any (reported per cycle)
      ■ Any (reported per patient)
    o Pregnancy Complications
      ■ Multiple births (and associated complications)
      ■ Ectopic pregnancies
      ■ Miscarriage
    o Neonatal Outcomes
      ■ Death
      ■ Birthweight
      ■ Congenital anomalies
    o Time to Pregnancy
      ■ Calendar time (months)
      ■ Number of cycles
    o Costs
      ■ Patient
      ■ Health system
      ■ Societal
    o Short-term Adverse Effects of Treatment
      ■ OHSS
• Surgical complications
  o Long-term Outcomes – Child
    ▪ Neurodevelopment / other issues related to prematurity
    ▪ Specific issues related to infertility treatment (epigenetic changes, sex chromosomal abnormalities, etc.)
    ▪ Cancer (all types)
  o Long-term Outcomes – Maternal
    ▪ Cancer
    ▪ Subsequent fertility
  o Donor Women Outcomes
    ▪ Short-term – OHSS
    ▪ Short-term – Surgical Complications
    ▪ Short-term – Adverse effects of treatments
    ▪ Long-term – Downstream fertility
    ▪ Long-term – Cancer
    ▪ Long-term – Age at menopause
    ▪ Quality of Life
  o Donor Men Outcomes
    ▪ Quality of life
    ▪ Short- and long-term health outcomes
  • Any additional description / clarification of the outcome reported on this form
  • Is this outcome form for a subgroup of interest? (Yes/No)
    o What subpopulation is this outcome reported for on this form?
      ▪ Age
      ▪ Race/ethnicity
      ▪ Obesity/BMI
      ▪ Ovarian reserve
      ▪ History of prior treatment
      ▪ Primary vs. secondary infertility
      ▪ Maternal parity
      ▪ Insurance status
      ▪ Diagnostic criteria / evaluation
      ▪ Presence or absence of male factor infertility
      ▪ Other female causes of infertility
      ▪ Hypertension
      ▪ Diabetes
      ▪ Women without male partners (single women or lesbian couples)
      ▪ Anatomic cause of tubal occlusion (e.g. prior sterilization vs. adhesions)
      ▪ Cause of male infertility
    o Any additional description / clarification of subgroup reported on this form
  • Total N Analyzed for this outcome
  • Timepoint reported on this form
    o Short-term
    o Long-term
  • Specify actual timing of the outcome (in months)
  • For each arm:
- N Analyzed (enter UNK if unknown)
- Unadjusted Result
  - Number of patients with outcome
  - % of patients with outcome
  - Events/denominator
  - Odds ratio
  - Hazard ratio
  - Relative risk
  - Mean
  - Median
  - Mean within group change
  - Mean between group change
  - Other (specify)
- Unadjusted Result Variability
  - 95% CI
  - IQR
  - Standard Error (SE)
  - Standard Deviation (SD)
  - Other % CI (specify)
  - Other (specify)
- Unadjusted Result, p-value between groups
- Unadjusted Result, indicate reference group (for comparison between groups)
- Adjusted Result
  - Number of patients with outcome
  - % of patients with outcome
  - Events/denominator
  - Odds ratio
  - Hazard ratio
  - Relative risk
  - Mean
  - Median
  - Mean within group change
  - Mean between group change
  - Other (specify)
- Adjusted Result Variability
  - 95% CI
  - IQR
  - Standard Error (SE)
  - Standard Deviation (SD)
  - Other % CI (specify)
  - Other (specify)
- Adjusted Result, p-value between groups
- Adjusted Result, indicate reference group (for comparison between groups)
- If adjusted data is recorded, indicate the adjustments applied
- Comments
Quality

- Study Type (select one): RCT, Cohort, Case-control, Cross-sectional
- If RCT, select Yes/No/Unclear for each of the following questions:
  - Selection Bias
    - Was the allocation sequence generated adequately (e.g., random number table, computer-generated randomization)?
    - Was the allocation of treatment adequately concealed (e.g., pharmacy-controlled randomization or use of sequentially numbered sealed envelopes)?
    - Were participants analyzed within the groups they were originally assigned to?
    - Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches?
  - Performance Bias
    - Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?
    - Did the study maintain fidelity to the intervention protocol?
  - Attrition Bias
    - If attrition (overall or differential nonresponse, dropout, loss to follow-up, or exclusion of participants) was a concern, were missing data handled appropriately (e.g., intention-to-treat analysis and imputation)?
  - Detection Bias
    - In prospective studies, was the length of follow-up different between the groups, or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls?
    - Were the outcome assessors blinded to the intervention or exposure status of participants?
    - Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants?
    - Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants?
  - Reporting Bias
    - Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported?
- If Cohort, select Yes/No/Unclear for each of the following questions:
  - Selection Bias
    - Were participants analyzed within the groups they were originally assigned to?
    - Did the study apply inclusion/exclusion criteria uniformly to all comparison groups?
    - Did the strategy for recruiting participants into the study differ across study groups?
    - Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches?
Performance Bias
- Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?
- Did the study maintain fidelity to the intervention protocol?

Attrition Bias
- If attrition (overall or differential nonresponse, dropout, loss to follow-up, or exclusion of participants) was a concern, were missing data handled appropriately (e.g., intention-to-treat analysis and imputation)?

Detection Bias
- In prospective studies, was the length of follow-up different between the groups, or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls?
- Were the outcome assessors blinded to the intervention or exposure status of participants?
- Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants?
- Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants?
- Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants?

Reporting Bias
- Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported?

If Case-Control, select Yes/No/Unclear for each of the following questions:

Selection Bias
- Were cases and controls selected appropriately (e.g., appropriate diagnostic criteria or definitions, equal application of exclusion criteria to case and controls, sampling not influenced by exposure status)
- Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches?

Performance Bias
- Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?
- Did the study maintain fidelity to the intervention protocol?

Attrition Bias
- If attrition (overall or differential nonresponse, dropout, loss to follow-up, or exclusion of participants) was a concern, were missing data handled appropriately (e.g., intention-to-treat analysis and imputation)?

Detection Bias
- In prospective studies, was the length of follow-up different between the groups, or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls?
- Were the outcome assessors blinded to the intervention or exposure status of participants?
Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants?
Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants?
Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants?

- Reporting Bias
  - Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported?

- If Cross-sectional, select Yes/No/Unclear for each of the following questions:
  - Selection Bias
    - Did the study apply inclusion/exclusion criteria uniformly to all comparison groups?
    - Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches?
  - Performance Bias
    - Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results?
  - Attrition Bias
    - If attrition (overall or differential nonresponse, dropout, loss to follow-up, or exclusion of participants) was a concern, were missing data handled appropriately (e.g., intention-to-treat analysis and imputation)?
  - Detection Bias
    - Were the outcome assessors blinded to the intervention or exposure status of participants?
    - Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants?
    - Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants?
    - Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants?
  - Reporting Bias
    - Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported?

- Other Bias
  - If applicable, describe any other concerns that may impact risk of bias

- Overall Study Rating (Good/Fair/Poor)
  - Good (low risk of bias). These studies have the least bias, and the results are considered valid. These studies adhere to the commonly held concepts of high quality, including the following: a clear description of the population, setting, approaches, and comparison groups; appropriate measurement of outcomes; appropriate statistical and analytical methods and reporting; no reporting errors; a low dropout rate; and clear reporting of dropouts.
  - Fair. These studies are susceptible to some bias, but not enough to invalidate the results. They do not meet all the criteria required for a rating of good quality.
because they have some deficiencies, but no flaw is likely to cause major bias. The study may be missing information, making it difficult to assess limitations and potential problems.

- **Poor** (high risk of bias). These studies have significant flaws that may have invalidated the results. They have serious errors in design, analysis, or reporting; large amounts of missing information; or discrepancies in reporting.

- If the study is rated as “Fair” or “Poor,” provide rationale.

- **Outcome-specific quality rating**
  - Do you think that any of the outcomes abstracted for this study should be assigned a quality rating DIFFERENT from the overall study rating? (No/Yes)
    - If you think any of the abstracted outcomes should have a quality rating different from the overall study, please provide the outcome(s), rating(s) and rationale(s).

**Applicability** – Use the PICOS format to identify specific issues, if any, that may limit the applicability of the study.

- **Population (P)**
  - Study population demographics not representative of intended population
  - Narrow or unrepresentative severity/stage/comorbidity
- **Intervention (I)**
  - Treatment protocol not representative of current practice
  - Change in standard of care
- **Comparator (C)**
  - Comparator not representative of current practice
- **Outcomes (O)**
  - Timing of outcome assessment
- **Setting (S)**
  - Standards or access to care vary from US setting
  - Specialty population or level of care
- **Comments**
Appendix C. List of Included Studies


Kar S and 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.


Maher MA, Sayyed TM and Elkhouly N. Cervical mucus removal prior to intrauterine insemination: a randomized trial. BJOG 2018;125(7):841-847. PMID: 29078018.


Morad AWA and Abdelhamid AA. Prospective randomized study for hydrotubation with or without lidocaine before intrauterine insemination in unexplained infertility. Middle East Fertility Society Journal 2012;17(4):250-255.


Appendix D. List of Excluded Studies

All studies listed below were reviewed in their full-text version and excluded for the reasons cited. Reasons for exclusion signify only the usefulness of the articles for this study and are not intended as criticisms of the articles.

Not a full publication or full text not available:


Alasmari N. Randomized trial on the effect of prewashing the insemination catheter on the pregnancy outcome. Fertility and Sterility 2016;106:e346.


Hurley EG and De Franco E. Influence of paternal age on perinatal outcomes in pregnancies achieved with assisted reproductive technologies. Fertility and Sterility 2016;106:e166.


Londra LC and Mumford SL. Birth weight in singletons after autologous fresh transfer according to the ovarian hyperstimulation protocol used. Fertility and Sterility 2016;106:e170.


Maslow BL, Griffin D, Benadiva CA, et al. Prospective double-blind randomized placebo controlled clinical trial comparing pregnancy rates after co-administration of low dose hcg at the time of GNRH-agonist trigger or 35 hours later, for the prevention of ohss. Fertility and Sterility 2016;106:e58.


Rodriguez-Purata J, Santistevan A, Sekhon L, et al. 1+ 1>2: A cost effectiveness analysis of single embryo transfer with PGS in two successive cycles vs a double embryo transfer with PGS in one. Fertility and Sterility 2016;106:e337.


Siristatidis CS, Bhattacharya S and Maheshwari A. In vitro maturation in sub fertile patients with polycystic ovarian syndrome undergoing assisted reproduction. Cochrane Database of Systematic Reviews 2007.


Volovsky M, Healey M, MacLachlan VB, et al. Intrauterine human chorionic gonadotropin (HCG) infusion prior to embryo transfer (ET) may be detrimental to pregnancy rate. Fertility and Sterility 2016;106:e52.

Wang AY. Increased rate of adverse neonatal outcomes among twins following assisted reproductive technology. Fertility and Sterility 2016;106:e174.


Not available in English:


Not original data from an RCT, SR/MA, or observational study with comparator:


Jungheim ES and Odibo AO. Fertility treatment in women with polycystic ovary syndrome: a decision analysis of different oral ovulation induction agents. Fertil Steril 2010;94(7):2659-64. PMID: 20451181.


Qin N, Chen Q, Hong Q, et al. Flexibility in starting ovarian stimulation at different phases of the menstrual cycle for treatment of infertile women with the use of in vitro fertilization or intracytoplasmic sperm injection. Fertil Steril 2016. PMID: 27114329.


**Observational study sample size less than 100 subjects:**


**Not a study population of interest:**


Lee WL, Chang WH, Wang KC, et al. The risk of epithelial ovarian cancer of women with endometriosis may be varied greatly if diagnostic criteria are different: A nationwide population-based cohort study. Medicine (United States) 2015;94(39):e1633.


No comparator of interest:


**No outcomes of interest:**


Abu Hashim H, Bazeed M and Abd Elaal I. Minimal stimulation or clomiphene citrate as first-line therapy in women with polycystic ovary syndrome: a randomized controlled trial. Gynecol Endocrinol 2012;28(2):87-90. PMID: 21770837.


Ketabchi AA. Intracytoplasmic sperm injection outcomes with freshly ejaculated sperms and testicular or epididymal sperm extraction in patients with idiopathic cryptozoospermia. Nephro-Urology Monthly 2016;8(6).


Naeem A, Amjad F and Memon AS. Comparison of letrozole versus clomiphene citrate on ovulation and achieving a successful pregnancy. Pakistan Journal of Medical and Health Sciences 2017;11(3):1143-1145


Rancourt RC, Harris HR and Michels KB. Methylation levels at imprinting control regions are not altered with ovulation induction or in vitro fertilization in a birth cohort. Hum Reprod 2012;27(7):2208-16. PMID: 22587996.


Outcomes not reported by underlying diagnosis or by using a multivariate model that includes diagnosis as one of the covariates:


Bellavia M, de Geyter C, Streuli I, et al. Randomized controlled trial comparing highly purified (HP-hCG) and recombinant hCG (r-hCG) for triggering ovulation in ART. Gynecol Endocrinol 2013;29(2):93-7. PMID: 23116325.


Cheung CS, Chan CH and Ng EH. Stress and anxiety-depression levels following first-trimester miscarriage: a comparison between women who conceived naturally and women who conceived with assisted reproduction. Bjo 2013;120(9):1090-7. PMID: 23631687.


Corchia C, Da Frè M, Di Lallo D, et al. Mortality and major morbidities in very preterm infants born from assisted conception or naturally conceived: Results of the area-based ACTION study. BMC Pregnancy and Childbirth 2014;14(1).


Dude AM, Yeh JS and Muasher SJ. Donor oocytes are associated with preterm birth when compared to fresh autologous in vitro fertilization cycles in singleton pregnancies. Fertility and Sterility 2016;106(3):660-665.


D-70


Kawahara S, Ueda A, Nakahori T, et al. Treatment period and medical care costs to achieve the first live birth by assisted reproductive technology are lower in the single embryo transfer period than in the double embryo transfer period: a retrospective analysis of women younger than 40 years of age. 2017;16(2):139-142.


Kwon H, Choi DH and Kim EK. Absolute position versus relative position in embryo transfer: A randomized controlled trial. Reproductive Biology and Endocrinology 2015;13(1).


Maimburg RD and Vaeth M. Do children born after assisted conception have less risk of developing infantile autism? Hum Reprod 2007;22(7):1841-3. PMID: 17456530.


Naether OGJ, Tandler-Schneider A and Bilger W. Individualized recombinant human follicle-stimulating hormone dosing using the CONSORT calculator in assisted reproductive technology:


2010 Society for Assisted Reproductive Technology Clinic Outcome Reporting System registry. Fertility and Sterility 2015.


Vuong TN, Phung HT and Ho MT. Recombinant follicle-stimulating hormone and recombinant luteinizing hormone versus recombinant follicle-stimulating hormone alone during GnRH antagonist ovarian stimulation in patients aged >/=35 years: a randomized controlled trial. Hum Reprod 2015;30(5):1188-95. PMID: 25740882.


**Does not meet study design criteria by outcome type:**


Elgafor El Sharkwy IA. Metformin versus laparoscopic unilateral ovarian drilling in clomiphene resistant women with polycystic ovary syndrome. Middle East Fertility Society Journal 2013;18(3):202-207.


Goudarzi ZM, Fallahzadeh H, Aflatoonian A, et al. Laparoscopic ovarian electrocautery versus gonadotropin therapy in infertile women with clomiphene citrate-resistant polycystic ovary


Takashima A, Takeshita N and Kinoshita T. Pregnancy outcomes after assisted reproductive procedures with embryos that had been derived from affected and unaffected ovaries among women with small unilateral endometriomas. Reprod Med Biol 2017;16(2):152-156. PMID: 29259463.


Appendix E. Characteristics of Included Studies

Table E-1 shows the study characteristics for the included studies. For full study citations, please refer to the report’s main reference list.

Table E-1. Characteristics of included studies

<table>
<thead>
<tr>
<th>Study ACRONYM</th>
<th>Companion Studies</th>
<th>Study Design</th>
<th>Geographic Location</th>
<th>N Enrolled</th>
<th>N Completed</th>
<th>Underlying Diagnosis</th>
<th>Interventions</th>
<th>Outcomes (Subgroups analyzed)</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdellah, 2011(^{156})</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Africa</td>
<td>147 140</td>
<td>PCOS</td>
<td></td>
<td>Letrozole 5 mg/d for 5 days, maximum treatment duration 6 consecutive cycles. vs. Laparoscopic ovarian drilling (LOD)</td>
<td>Live birth Miscarriage Multiple births</td>
<td>Good</td>
</tr>
<tr>
<td>Aboulgahar, 2010(^{162})</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Africa</td>
<td>84 NR</td>
<td>PCOS</td>
<td></td>
<td>Routine IVF/ICSI using highly purified uFSH (Fostimon) vs. rFSH (Gonal F)</td>
<td>OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Abu Hashim, 2010(^{168})</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Africa</td>
<td>260 260</td>
<td>PCOS</td>
<td></td>
<td>Letrozole, 2.5 mg/d for 5 days, maximum treatment duration 6 cycles vs. Laparoscopic ovarian diathermy (LOD)</td>
<td>Live birth Miscarriage Multiple births OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Abu Hashim, 2011(^{154})</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Africa</td>
<td>176 165</td>
<td>PCOS</td>
<td></td>
<td>Laparoscopic ovarian diathermy (LOD) performed at least 8 weeks after the last CC dosage vs. CC 50-150 mg/d for 5 days, maximum treatment duration 6 cycles</td>
<td>Live birth Miscarriage OHSS</td>
<td>Fair</td>
</tr>
<tr>
<td>Abu Hashim, 2011(^{157})</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Middle East</td>
<td>188 188</td>
<td>PCOS</td>
<td></td>
<td>CC 50 - 150 mg/d, maximum treatment duration 3 cycles followed by IUI vs. Timed intercourse</td>
<td>Live birth Ectopic pregnancy Miscarriage Multiple births OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Abu Hashim, 2012(^{204})</td>
<td>KQ 2</td>
<td>RCT</td>
<td>Africa</td>
<td>136 125</td>
<td>Endometriosis</td>
<td></td>
<td>IUI following hCG injection using Letrozole 5 mg/d vs. CC 100 mg/d on cycle days 3-9</td>
<td>Live birth Miscarriage</td>
<td>Good</td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled</td>
<td>N Completed</td>
<td>Underlying Diagnosis (Subgroups analyzed)</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>----------------------------------</td>
<td>------------</td>
<td>-------------</td>
<td>------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Aghahosseini, 2017</td>
<td>KQ 1</td>
<td>RCT Middle East</td>
<td>100 80</td>
<td>100 80</td>
<td>PCOS</td>
<td>Low dose hCG 35 hours after GnRH agonist</td>
<td>OHSS</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Amer, 2009</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>72 65</td>
<td>72 65</td>
<td>PCOS</td>
<td>Laparoscopic ovarian diathermy for endometriosis, periadnexal adhesions, and adhesions</td>
<td>Miscarriage</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Amer, 2017</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>159 149</td>
<td>159 149</td>
<td>PCOS</td>
<td>Clomiphene vs. Letrozole</td>
<td>Live birth</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>An, 2014</td>
<td>KQ 1</td>
<td>RCT Asia</td>
<td>150 109</td>
<td>150 109</td>
<td>PCOS</td>
<td>Berberine was administered at a dosage of 3 x 500 mg daily for greater than or equal to 12 weeks before controlled ovarian stimulation. vs. Metformin was administered at a dosage of 3 x 500 mg daily for greater than or equal to 12 weeks before controlled ovarian stimulation. vs. Placebo was administered as one tablet three times daily for greater than or equal to 12 weeks before controlled ovarian stimulation.</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Badawy, 2008</td>
<td>KQ 1</td>
<td>RCT Middle East</td>
<td>318 318</td>
<td>318 318</td>
<td>PCOS</td>
<td>Clomiphene citrate vs. Gonadotrophin</td>
<td>Miscarriage</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Badawy, 2009</td>
<td>KQ 3</td>
<td>RCT Middle East</td>
<td>996 996</td>
<td>996 996</td>
<td>Unknown</td>
<td>Letrozole 5 mg/d for 5 days vs. Anastrozole 1 mg/d for 5 days vs. CC 100 mg/d for 5 days. vs. Spontaneous pregnancy</td>
<td>Miscarriage</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bagis, 2010&lt;sup&gt;221&lt;/sup&gt; KQ 3, 5</td>
<td>RCT UK/Europe</td>
<td>228 226 Unknown, Male</td>
<td>IUI performed 36 hours after hCG injection vs. IUI performed 18 hours after hCG followed by second IUI performed 40 hours after hCG</td>
<td>Live birth (male) Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balaban, 2011&lt;sup&gt;258&lt;/sup&gt; KQ 5</td>
<td>RCT Middle East</td>
<td>77 cycles 77 cycles Male</td>
<td>ICSI vs. IMSI</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barad, 2017&lt;sup&gt;280&lt;/sup&gt; Across All KQs</td>
<td>Observational US</td>
<td>33,756 21,008 All</td>
<td>PGD vs. Non-PGD in donor oocyte cycles</td>
<td>Live birth Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belva, 2011&lt;sup&gt;259&lt;/sup&gt; KQ 5</td>
<td>Observational UK/Europe</td>
<td>120 120 Male</td>
<td>Male offspring born to parents who underwent ICSI vs. Male offspring born to parents who conceived spontaneously</td>
<td>Birthweight</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bhattacharya, 2008&lt;sup&gt;115&lt;/sup&gt; SUIT KQ 3</td>
<td>RCT UK/Europe</td>
<td>580 576 Unknown</td>
<td>Expectant Management vs. CC 50 mg days 2-6 of cycle vs. IUI</td>
<td>Live birth (diagnostic criteria) Time to pregnancy Ectopic pregnancy Miscarriage Patient costs</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bodri, 2008&lt;sup&gt;268&lt;/sup&gt; KQ 6</td>
<td>Observational UK/Europe</td>
<td>2,653 2,653 Donor</td>
<td>Ovarian stimulation with GnRH agonist vs. Ovarian stimulation with GnRH antagonist/hCG vs. Ovarian stimulation with GnRH antagonist/GnRH agonist vs. A control group was created by taking into account all IVF cycles reaching oocyte retrieval performed during the same period</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bodri, 2009&lt;sup&gt;269&lt;/sup&gt; KQ 6</td>
<td>Observational UK/Europe</td>
<td>1,171 1,171 Donor</td>
<td>Triggering with recombinant hCG vs. Triggering with GnRH</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>---------------------------------</td>
<td>------------------------------------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boulet, 2015&lt;sup&gt;254&lt;/sup&gt; NASS KQ 5</td>
<td>Observational US</td>
<td>499,135 cycles NA Male</td>
<td>Conventional IVF vs. ICSI</td>
<td>Live birth Miscarriage Multiple pregnancies Birthweight</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boulet, 2016&lt;sup&gt;277&lt;/sup&gt; Across All KQs</td>
<td>Observational US</td>
<td>4,618,076 4,618,076 All</td>
<td>No intervention (spontaneous conception) vs. Conventional IVF vs. IVF + ICSI</td>
<td>Congenital anomalies</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brinton, 2015&lt;sup&gt;136&lt;/sup&gt; Across All KQs</td>
<td>Observational US</td>
<td>9,892 9,892 All</td>
<td>Control vs. CC vs. Gonadotropins</td>
<td>Maternal cancer</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Butts, 2014&lt;sup&gt;122&lt;/sup&gt; NASS KQ 3</td>
<td>Observational US</td>
<td>38,926 38,926 Unknown</td>
<td>ICSI vs. IVF vs. Assisted hatching vs. No assisted hatching</td>
<td>Live birth (ovarian reserve)</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chang, 2016&lt;sup&gt;276&lt;/sup&gt; Across All KQs</td>
<td>Observational US</td>
<td>106,902 cycles 106,902 cycles All</td>
<td>PGD vs. PGD done for concern of aneuploidy vs. PGD done for other concern outside of genetics/aneuploidy vs. No PGD</td>
<td>Live birth Multiple births Miscarriage Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chen, 2016 KQ 1</td>
<td>RCT Asia</td>
<td>1,508 1,508 PCOS</td>
<td>Frozen embryo transfer vs. Fresh embryo transfer</td>
<td>Live birth, Miscarriage, Multiple births, Birthweight, Birthweight, OHSS, Ectopic pregnancy, Neonatal death, Congenital anomalies</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Choi, 2012 KQ 1</td>
<td>RCT Asia</td>
<td>61 61 PCOS</td>
<td>IVM/IVF with FSH and hCG priming protocol vs. GnRH agonist long protocol group vs. GnRH antagonist multi-dose flexible protocol</td>
<td>Live birth</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crawford, 2017 Across All KQs</td>
<td>Observational US</td>
<td>105,517 cycles 105,517 cycles All</td>
<td>Autologous cycles: Cryopreserved oocyte vs. Fresh oocyte Donor cycles: Cryopreserved oocyte vs. Fresh oocyte</td>
<td>Live birth, Miscarriage</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custers, 2012 KQ 3</td>
<td>RCT UK/Europe</td>
<td>253 253 Unknown</td>
<td>Expectant Management for 6 months, followed by six cycles of IUI-COS, followed by 3 cycles of IVF vs. IUI with controlled ovarian stimulation (IUI-COS) for 6 months followed by 3 cycles of IVF</td>
<td>Miscarriage, Ectopic pregnancy, Multiple births, Health system costs</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Danhof, 2018 KQ 3</td>
<td>RCT UK/Europe</td>
<td>738 738 Unknown</td>
<td>CC with IUI vs. FSH with IUI</td>
<td>Live birth, Ectopic pregnancy, Miscarriage, Multiple births, OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>--------------------------------</td>
<td>------------------------------------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>de Wilde, 2017(^{280}) KQ 1</td>
<td>Observational UK/Europe</td>
<td>3,077 3,077 PCOS</td>
<td>Natural conception vs. Ovulation Induction vs. IVF-ICSI</td>
<td>Birthweight Neonatal death</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Demirol, 2007\(^{210}\) KQ 3 | RCT UK/Europe | 241 241 Unknown | IUI following:
  - Group I (Gonal-F, Serono, Turkey), 81 Follitropin alpha vs.
  - Group II (Metrodin-HP, Serono), highly-purified uFSH vs.
  - Group III (Pergonal, Serono), hMG | Miscarriage OHSS | Good |
<p>| Dhalwani, 2016(^{236}) KQ 3 and 5 | Observational US | 3,896,242 births 3,896,242 births Unknown, Male | ART vs. No intervention (spontaneous conception) | Birthweight | Good |
| Diamond, 2015(^{233}) KQ 3 | RCT US | 900 746 Unknown | Gonadotropin vs. Clomiphene vs. Letrozole | Live birth Multiple births Miscarriage Ectopic pregnancy Birthweight Neonatal death Congenital anomalies | Good |
| Dreyer, 2016(^{250}) KQ 4 | RCT UK/Europe | 85 85 (for ITT analysis) Tubal | Hysteroscopic proximal occlusion by intratubal device placement (Essure Device) vs. Laparoscopic salpingectomy | Live birth Miscarriage Ectopic pregnancy Time to pregnancy | Good |
| Ebrahimi, 2010(^{229}) KQ 3 | RCT Middle East | 200 179 Unknown | CC 50 mg BID on cycle days 3-7 followed by 75 IU hMG on cycle days 7-9 and adjusted thereafter. IUI performed following triggered ovulation. Cyclogest vaginal pessaries 400 mg/daily through the tenth week of pregnancy vs. No luteal phase support | Live birth Multiple births | Good |</p>
<table>
<thead>
<tr>
<th>Study ACRONYM KQs</th>
<th>Study Design Geographic Location</th>
<th>N Enrolled N Completed Underlying Diagnosis</th>
<th>Interventions</th>
<th>Outcomes (Subgroups analyzed)</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Einarsson, 2017¹</td>
<td>RCT UK/Europe</td>
<td>962 317 PCOS</td>
<td>Lifestyle change (weight reduction) +IVF vs. IVF only</td>
<td>Live birth Multiple births Miscarriage Ectopic pregnancy OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Elsedeek, 2014¹</td>
<td>RCT Middle East</td>
<td>220 cycles 220 cycles PCOS</td>
<td>Clomiphene citrate 200mg/day over 5 days vs. Clomiphene citrate 100mg/day over 10 days</td>
<td>Live birth</td>
<td>Fair</td>
</tr>
<tr>
<td>Emekci, 2017²</td>
<td>RCT Middle East</td>
<td>196 196 PCOS</td>
<td>4g Myo-Inositol (MYO) plus 400 mg folic acid vs. Recombinant FSH and no MYO administration</td>
<td>Miscarriage</td>
<td>Good</td>
</tr>
<tr>
<td>Erdem, 2009³</td>
<td>RCT UK/Europe</td>
<td>214 214 Unknown</td>
<td>Gonadotropin IUI followed by luteal support with crinone once a day beginning 2 days after insemination until pregnancy testing. vs. Gonadotropin IUI without luteal support.</td>
<td>Live birth</td>
<td>Fair</td>
</tr>
<tr>
<td>Erdem, 2015⁴</td>
<td>RCT Middle East</td>
<td>219 174 Unknown, Male</td>
<td>rFSH followed by triggered ovulation vs. CC 100 mg/d on days 3-7 of cycle followed by triggered ovulation</td>
<td>Live birth</td>
<td>Good</td>
</tr>
<tr>
<td>Farquhar, 2018⁵</td>
<td>RCT Australia/N.Z.</td>
<td>473 201 Unknown</td>
<td>IUI vs. Expectant Management</td>
<td>Live birth Multiple births Miscarriage Ectopic pregnancy</td>
<td>Good</td>
</tr>
<tr>
<td>Ge, 2008⁶</td>
<td>RCT Asia</td>
<td>62 62 PCOS</td>
<td>IVF using oocytes cultured in media containing hCG, rFSH, and rhCG vs. hCG-free media with rFSH for the first 10 hours, then were transferred to the same medium the group above vs. hCG-free media only</td>
<td>Live birth Miscarriage</td>
<td>Good</td>
</tr>
<tr>
<td>Study ACRONYM KQs</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
</tr>
<tr>
<td>-------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Ghahiri, 2016170</td>
<td>RCT Middle East</td>
<td>101 101 PCOS</td>
<td>Clomiphene 100 mg vs. Letrozole 5 mg</td>
<td>Multiple births Miscarriage Ectopic pregnancy OHSS Time to pregnancy</td>
<td>Fair</td>
</tr>
<tr>
<td>Ghanem, 2013145</td>
<td>RCT Middle East</td>
<td>174 159 PCOS</td>
<td>CC 100 mg/d for 5 days plus uFSH 37.5 IU/d. vs. uFSH 37.5 IU/d only</td>
<td>Live birth</td>
<td>Good</td>
</tr>
<tr>
<td>Gibreel, 2013228</td>
<td>RCT Africa</td>
<td>105 90 Unknown</td>
<td>Endometrial scratching using a pipelle biopsy catheter with biopsies obtained vs. Sham procedure using uterine sound only</td>
<td>Miscarriage Multiple births</td>
<td>Good</td>
</tr>
<tr>
<td>Goldman, 2014210</td>
<td>RCT US</td>
<td>154 115 Unknown</td>
<td>CC 100 mg/d for 5 days followed by IUI, maximum 2 cycles after which patients proceeded to IVF up to 6 cycles vs. rFSH followed by IUI. maximum 2 cycles after which patients proceeded to IVF up to 6 cycles vs. Immediate IVF up to 6 cycles.</td>
<td>Live birth Time to pregnancy Ectopic pregnancy Miscarriage Multiple births Birthweight Birthweight Neonatal death OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Gregoriou, 2008236</td>
<td>RCT UK/Europe</td>
<td>50 50 Unknown</td>
<td>rFSH beginning on cycle day 3 vs. Letrozole 5 mg/d on cycle day 3</td>
<td>Live birth (prior treatments)</td>
<td>Good</td>
</tr>
<tr>
<td>Grimstad, 2016251</td>
<td>Observational US</td>
<td>7,145 7,145 Tubal</td>
<td>IVF vs. ICSI</td>
<td>Live birth Multiple births Miscarriage Birthweight</td>
<td>Good</td>
</tr>
<tr>
<td>Hajizadeh, 2017261</td>
<td>RCT Middle East</td>
<td>419 386 Male</td>
<td>Structured aerobic exercise for 12 weeks vs. No exercise</td>
<td>Live birth</td>
<td>Good</td>
</tr>
<tr>
<td>Harira, 2018246</td>
<td>RCT Middle East</td>
<td>172 172 Unknown</td>
<td>CC + estradiol vs. Letrozole</td>
<td>Ectopic pregnancy Miscarriage OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>Hassan, 2017193</td>
<td>RCT Middle East</td>
<td>182 140 PCOS</td>
<td>Letrozole vs. FSH</td>
<td>Miscarriage Adverse events Cost effectiveness</td>
<td>Good</td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design</td>
<td>N Enrolled</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>------------</td>
<td>---------------</td>
<td>------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>KQs Companion Studies</td>
<td>Geographic Location</td>
<td>N Enrolled</td>
<td>Underlying Diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hershko-Klement, 2016</td>
<td>Observational Middle East</td>
<td>2,406</td>
<td>ICSI vs. IMSI</td>
<td>Congenital anomalies</td>
<td>Fair</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1,981</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homburg, 2012</td>
<td>RCT UK/Europe &amp; Latin America</td>
<td>302</td>
<td>CC 50 – 150 mg/d for 5 days, triggered ovulation vs. rhFSH, triggered ovulation</td>
<td>Live birth Ectopic pregnancy Miscarriage Multiple births</td>
<td>Good</td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td>255</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosseini, 2010</td>
<td>RCT Middle East</td>
<td>NR</td>
<td>Long-term desensitization protocol using GnRH agonist buserelin 500 mcg SQ. Gonal F started on day 3, replaced by hMG after 7th day of stimulation. vs. Gonal F for ovarian stimulation. Cetrorelix (GnRH antagonist) 0.35 mg/d injected SQ for 3 days. hMG prescribed after 7th day of stimulation. Fertilization via ICSI. 3 good quality embryos transferred 3 days later.</td>
<td>Miscarriage OHSS</td>
<td>Fair</td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td>112</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosseini-Rashidi, 2016</td>
<td>RCT Middle East</td>
<td>104</td>
<td>Clomiphene vs. Recombinant human fSH</td>
<td>Live birth Miscarriage OHSS</td>
<td>Fair</td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td>96</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ibrahim, 2017</td>
<td>RCT Middle East</td>
<td>80</td>
<td>Laparoscopic Ovarian Drilling vs. Letrozole</td>
<td>Miscarriage</td>
<td>Fair</td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td>80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jacob, 2016</td>
<td>RCT UK/Europe</td>
<td>153</td>
<td>Placebo vs. Metformin</td>
<td>Live birth OHSS</td>
<td>Good</td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td>153</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PCOS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jahromi, 2017</td>
<td>RCT Middle East</td>
<td>80</td>
<td>Melatonin along with ART vs. Placebo with ART</td>
<td>Miscarriage</td>
<td>Fair</td>
</tr>
<tr>
<td>KQ 3</td>
<td></td>
<td>66</td>
<td>Unknown</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

E-9
<table>
<thead>
<tr>
<th>Study ACRONYM</th>
<th>KQs Companion Studies</th>
<th>Study Design</th>
<th>Geographic Location</th>
<th>N Enrolled</th>
<th>N Completed</th>
<th>Underlying Diagnosis</th>
<th>Interventions</th>
<th>Outcomes (Subgroups analyzed)</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kansal Kalra, 2008</td>
<td>KQ 3</td>
<td>RCT</td>
<td>US</td>
<td>223</td>
<td>18</td>
<td>Unknown</td>
<td>Follicular arm - rFSH on cycle day 1 or 2 of the oocyte retrieval cycle. vs. Luteal phase arm - rFSH 9 days after spontaneous LH surge of the menstrual cycle preceding oocyte retrieval</td>
<td>Live birth (ovarian reserve)</td>
<td>Fair</td>
</tr>
<tr>
<td>Kar, 2015</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Asia</td>
<td>173</td>
<td>80</td>
<td>PCOS</td>
<td>Clomiphene vs. Metformin vs. Clomiphene + Metformin</td>
<td>Live birth Miscarriage</td>
<td>Fair</td>
</tr>
<tr>
<td>Kettner, 2016</td>
<td>KQ 1, 4, 5</td>
<td>Observational</td>
<td>UK/Europe</td>
<td>184</td>
<td>565,116</td>
<td>PCOS, Tubal, Male</td>
<td>No fertility treatment vs. IUI or OI vs. IVF or ICSI</td>
<td>Type 1 diabetes in children</td>
<td>Good</td>
</tr>
<tr>
<td>Keyhan, 2018</td>
<td>KQ 1, 2, 3, 4, 5</td>
<td>Observational</td>
<td>US</td>
<td>185</td>
<td>90,401</td>
<td>PCOS, Endometriosis, Unknown, Tubal, Male</td>
<td>ICSI vs. IVF</td>
<td>Birthweight (male infertility, age)</td>
<td>Fair</td>
</tr>
<tr>
<td>Khosravi, 2015</td>
<td>KQ 3</td>
<td>RCT</td>
<td>Middle East</td>
<td>232</td>
<td>180</td>
<td>Unknown</td>
<td>Oral dydrogesterone vs. Vaginal cyclogest</td>
<td>Miscarriage</td>
<td>Fair</td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td>--------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, 2011&lt;sup&gt;210&lt;/sup&gt;</td>
<td>RCT Asia</td>
<td>110 110 Unknown</td>
<td>IVF/ICSI with testosterone gel pretreatment (12.5 mg/d) starting on cycle day 6 of the estrogen-progesterone pretreatment. vs. 21 days pretreatment with estradiol valerate and norethindrone</td>
<td>Live birth (ovarian reserve) Miscarriage (ovarian reserve)</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kim, 2012&lt;sup&gt;150&lt;/sup&gt;</td>
<td>RCT Asia</td>
<td>211 208 PCOS</td>
<td>Ovarian stimulation using 50 - 150 IU of rhFSH after establishing ovarian and uterine quiescence using vaginal ultrasound. GnRH antagonist, cetrorelix (Cetrotide) 0.125 mg/d was administered in the morning of stimulation day 1 and 2. When the mean diameter of lead follicle reached 13 mm, cetrorelix at a dose of 0.25 mg/d was started again and continued daily up to the day of rhCG (injection). vs. GnRH agonist, triptorelin (Decapeptyl) at a dose of 0.1 mg/d was initiated from day 18 of oral contraceptive pretreatment cycle. All patients had withdrawal bleeding after discontinuation of oral contraceptive. When pituitary desensitization was achieved, ovarian stimulation was started and the dose of triptorelin was reduced to 0.05 mg daily and continues up to day of rhCG administration. Ovarian stimulation was performed in the same manner.</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kissin, 2015&lt;sup&gt;272&lt;/sup&gt;</td>
<td>Observational US</td>
<td>42,383 42,383 All</td>
<td>Singleton ICSI vs. Singleton Conventional IVF (without ICSI) vs. Multiples ICSI vs. Multiples Conventional IVF (without ICSI)</td>
<td>Neurodevelopment-autism diagnosis (male)</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Across All KQs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td>---------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kjotrod, 2011</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>150 149 PCOS</td>
<td>Metformin prior to, and during, AFT vs. Placebo prior to, and during, AFT</td>
<td>Live birth</td>
<td>Fair (ITT results) Poor (non-ITT results)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Knudtson, 2017</td>
<td>Across All KQs</td>
<td>Observational US</td>
<td>151,533 first-cycle 151,533 first-cycle All</td>
<td>Frozen embryo without assisted hatching vs. Frozen embryo with assisted hatching</td>
<td>Live birth (race, age, etiology of infertility)</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kramer, 2009</td>
<td>KQ 6</td>
<td>Observational US</td>
<td>287 155 Donor</td>
<td>All respondents were egg donors vs. No comparator</td>
<td>Adverse effects of treatments OHSS Downstream fertility</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurzawa, 2008</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>74 70 PCOS</td>
<td>All patients received oral contraceptives pills x 1 month before starting controlled ovarian hyperstimulation. None of the patients used oral anti diabetic medications (biguanides or thiazolidinediones). rhFSH started on cycle day 2 at 150 IU/d and adjusted depending on an ovarian response. A GnRH antagonist - cetrorelix 0.25 mg subcutaneous injections were given until the criteria for recombinant hCG administration were met vs. During oral contraception on days 16–18 of the preceding cycle, after transvaginal ultrasonographic screening of ovaries, an intramuscular injection of GnRH agonist triptorelin (Diphereline SR 3.75; Boufor Ibsen Pharma, France) was given. After confirmation of pituitary desensitization (LH &lt;2 mIU/mL and estradiol &lt;40 pg/mL) the administration of FSH was commenced. rFSH and hCG administered as above</td>
<td>Live birth Miscarriage Multiple births OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kuzmin, 2014248 KQ 4</td>
<td>RCT Asia</td>
<td>468 468 Tubal Factor</td>
<td>Laparoscopy, salpingolysis, salpingostomy, and transcervical falloposcopy tubal dilatation (TFTD) vs. Laparoscopy, salpingolysis, salpingostomy</td>
<td>Ectopic pregnancy</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>La Sala, 2015243 KQ 5</td>
<td>RCT UK/Europe</td>
<td>242 242 Male</td>
<td>IMSI vs. ICSI</td>
<td>Live birth Miscarriage Multiple births Birthweight Congenital anomalies</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leandri, 2013256 KQ 5</td>
<td>RCT UK/Europe</td>
<td>255 255 Male</td>
<td>Conventional ICSI vs. ICSI</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2007128 PPCOS KQ 1 Companion: Rausch, 2009129</td>
<td>RCT US</td>
<td>626 450 PCOS</td>
<td>CC initial dose of 50 mg of CC on days 3–7 of each treatment cycle on-study vs. Metformin 500 mg/d, increased to 2000 mg/d Vs. Combination of CC and metformin</td>
<td>Live birth (obesity/BMI) Ectopic pregnancy Miscarriage Multiple births Congenital anomalies</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2014131 PPCOS 2 KQ 1 Companion: Polotsky, 2015132</td>
<td>RCT US</td>
<td>750 750 PCOS</td>
<td>CC 50 mg daily starting on cycle day 3 for 5 days vs. Letrozole 2.5 mg daily starting on cycle day 3 for 5 days Maximum treatment duration up to 5 cycles</td>
<td>Live birth Time to pregnancy Ectopic pregnancy Miscarriage Birthweight Congenital anomalies Neonatal death</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Legro, 2015172 KQ 1</td>
<td>RCT US</td>
<td>149 132 PCOS</td>
<td>Continuous OCP vs. Lifestyle changes (caloric restriction, physical activity, weight loss medication) vs. Combined OCP and lifestyle changes</td>
<td>Live birth Birthweight Ectopic pregnancy Miscarriage</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Levi Dunietz, 2017⁷⁻²⁹⁷ KQ 3</td>
<td>Observational US</td>
<td>4,292,779 4,292,779 Unknown</td>
<td>Non-ART vs. Fresh embryos vs. Cryopreserved embryos</td>
<td>Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Litzky, 2018⁴⁴ Across All KQs</td>
<td>Observational US</td>
<td>1,008,393 180,184 All</td>
<td>Fresh embryo transfers vs. Frozen/thawed embryo transfers</td>
<td>Birthweight</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Litzky, 2018⁷⁹⁰ Across All KQs</td>
<td>Observational US</td>
<td>124,154 124,154 All</td>
<td>Blastocyst stage transfer vs. Cleavage stage transfer</td>
<td>Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Londra, 2016³⁷⁸ Across All KQs</td>
<td>Observational US</td>
<td>136,605 cycles that resulted in a pregnancy 136,605 All</td>
<td>Luteal GnRH agonist cycles vs. GnRH antagonist vs. GnRH agonist flare cycles</td>
<td>Ectopic pregnancy</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luke, 2010¹³³ NASS KQ 2, 3, 4, 5</td>
<td>Observational US</td>
<td>69,028 cycles 69,028 cycles Endometriosis, Unknown, Tubal Factor, Male</td>
<td>Elective single embryo transfer (eSET)→1 embryo vs. eSET→2 embryos vs. eSET→3 embryos vs. eSET→4 or more embryos</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luke, 2016³⁷⁴ Across All KQs</td>
<td>Observational US</td>
<td>53,859 53,859 All</td>
<td>ART vs. ART</td>
<td>Birthweight Congenital anomalies</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----------------------------------</td>
<td>------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Magnusson, 2018&lt;sup&gt;287&lt;/sup&gt; Across All KQs</td>
<td>Observational UK/Europe</td>
<td>44,369 27,359 287 All</td>
<td>Number of oocytes retrieved: &lt;10 vs. 10-14 vs. 15-19 vs. &gt;20</td>
<td>Birthweight Congenital anomalies Neonatal death</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maher, 2018&lt;sup&gt;245&lt;/sup&gt; KQ 3</td>
<td>RCT Middle East</td>
<td>714 714 Unknown</td>
<td>Unstimulated IUI: Cervical mucus removal (internal and external) vs. No mucus removal</td>
<td>Live birth Miscarriage Multiple births</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Majumdar, 2013&lt;sup&gt;214&lt;/sup&gt; KQ 3</td>
<td>RCT Asia</td>
<td>156 151 Unknown</td>
<td>ICSI with sperm selection based on visual assessment vs. ICSI with sperm selection based on ability to bind hyaluronic acid</td>
<td>Live birth Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malchau, 2017&lt;sup&gt;196&lt;/sup&gt; KQ 1, 2, 3, 5</td>
<td>Observational UK/Europe</td>
<td>19,884 19,884 PCOS, Endometriosis, Unknown, Male</td>
<td>IUI vs. ART vs. No treatment</td>
<td>Live birth Time to birth</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mancuso, 2016&lt;sup&gt;233&lt;/sup&gt; Across All KQs</td>
<td>Observational US</td>
<td>914 914 All</td>
<td>Elective single embryo transfer (eSET) vs. Double embryo transfer (DET)</td>
<td>Live birth Multiple birth</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maxwell, 2008&lt;sup&gt;770&lt;/sup&gt; KQ 6</td>
<td>Observational US</td>
<td>587 587 Donor</td>
<td>Oocyte Donors vs. No comparison</td>
<td>Adverse effects of treatments OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mehrabian, 2012&lt;sup&gt;148&lt;/sup&gt; KQ 1</td>
<td>RCT Middle East</td>
<td>104 Unclear PCOS</td>
<td>hMG followed by triggered ovulation vs. Laparoscopic ovarian drilling</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mohammadi, 2018&lt;sup&gt;190&lt;/sup&gt;</td>
<td>KQ 1</td>
<td>RCT Middle East</td>
<td>219 184 PCOS</td>
<td>Methylprednisolone vs. No treatment</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morad, 2012&lt;sup&gt;227&lt;/sup&gt;</td>
<td>KQ 3</td>
<td>RCT Africa</td>
<td>234 231 Unknown</td>
<td>Hydrotubation performed one day before IUI using: 20 mL of saline vs. 20 mL of 0.1 mg Lidocaine/mL saline mixed with 19.9 cc of saline</td>
<td>Ectopic pregnancy Miscarriage Multiple births OHSS</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morin-Papunen, 2012&lt;sup&gt;151&lt;/sup&gt;</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>320 259 PCOS</td>
<td>Metformin (500 mg) was initiated at a dose of one tablet once a day for the first week and increased thereafter by one tablet daily in weekly steps up to three tablets (one + two daily) in nonobese women and to four tablets (two + two daily) in obese women and was continued up to a maximum of 9 months. If pregnancy occurred, metformin was continued up to the 12th week. The women used metformin or placebo alone for at least 3 months. If pregnancy did not occur, ovulation induction was commenced: if the woman ovulated after CC, she continued metformin/placebo with the same dose of clomiphene for four to six cycles or until the 12th week of pregnancy. After four to six unsuccessful cycles with metformin/placebo and CC, either gonadotrophins or aromatase inhibitors were used. vs. Placebo with all other management the same as above.</td>
<td>Live birth (obesity/BMI) Time to pregnancy</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs</td>
<td>Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed</td>
<td>Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>------------------</td>
<td>---------------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Muller, 2017</td>
<td>KQ 2</td>
<td></td>
<td>RCT UK/Europe</td>
<td>144</td>
<td>144</td>
<td>Endometriosis</td>
<td>Dienogest vs. a-GnRH vs. No treatment</td>
<td>Live birth</td>
<td>Fair</td>
</tr>
<tr>
<td>Nada, 2016</td>
<td>KQ 3</td>
<td></td>
<td>RCT Africa</td>
<td>622</td>
<td>595</td>
<td>Unknown</td>
<td>Pts given human menopausal gonadotropins form Day 2 to reach DF of 18-22 mm. Then Ganirelix acetate 0.25 mg SQ started from Day 6 or 7. Then IUI of 0.5 mL vs. Clomiphene citrate 100 mg from Day 2 to 6. Monitor ovulation, the hCG given at dose of 10,000 IU IM. Then, IUI of 0.5 mL</td>
<td>OHSS, Cost</td>
<td>Good</td>
</tr>
<tr>
<td>Nahuis, 2011</td>
<td>KQ 1</td>
<td>Companion: Nahuis, 2012</td>
<td>RCT UK/Europe</td>
<td>168</td>
<td>168</td>
<td>PCOS</td>
<td>Laparoscopic electrocautery (LEC). LEC followed by clomiphene citrate and then rFSH if still anovulatory. vs. rFSH</td>
<td>Live birth, Patient costs</td>
<td>Fair</td>
</tr>
<tr>
<td>Nandi, 2017</td>
<td>KQ 3</td>
<td></td>
<td>RCT UK/Europe</td>
<td>207</td>
<td>207</td>
<td>Unknown</td>
<td>Three cycles of IUI+COH vs. One cycle of IVF</td>
<td>Live birth, Multiple birth, Miscarriage, Ectopic pregnancy, OHSS</td>
<td>Fair</td>
</tr>
<tr>
<td>Nangia, 2011</td>
<td>NASS</td>
<td>KQ 5</td>
<td>Observational US</td>
<td>77,432 cycles</td>
<td>77,432 cycles</td>
<td>Male</td>
<td>IVF w/o ICSI, Male only vs. ICSI, Male only</td>
<td>Live birth, Birthweight</td>
<td>Fair</td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>------------------------</td>
<td>----------------------------------</td>
<td>---------------------------------------------</td>
<td>---------------</td>
<td>------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oyesanya 2009</td>
<td>KQ 3</td>
<td>RCT UK/Europe</td>
<td>353 351 Unknown</td>
<td>Provided that 6 or more oocytes were retrieved from the prospective donor, half were given to the recipient and half were given to another recipient. vs. Recipients received all retrieved oocytes from their altruistic donor.</td>
<td>Ectopic pregnancy (diagnostic criteria)</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palomba, 2010</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>50 47 PCOS</td>
<td>Laparoscopic ovarian diathermy. No drugs to trigger ovulation. vs. CC for up to 6 cycles plus metformin 500 mg tapered upwards. No drugs to trigger ovulation.</td>
<td>Live birth Surgical complications</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Palomba, 2011</td>
<td>KQ 1</td>
<td>RCT UK/Europe</td>
<td>120 120 PCOS</td>
<td>Metformin 500 mg three times daily vs. Placebo Metformin and placebo treatments started on the same day of GnRH-a administration (pretreatment) and were continued during the gonadotropin ovarian stimulation (cotreatment), treatment that started at least 14 days later. Both active drug and placebo were stopped when a positive pregnancy test or menstrual bleeding appeared.</td>
<td>Live birth OHSS</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Peeraer, 2015</td>
<td>Across All KQs</td>
<td>RCT UK/Europe</td>
<td>579 cycles 434 cycles All</td>
<td>Frozen/thawed embryo transfer in women with a natural cycle vs. Frozen/thawed embryo transfer in women injected with hMG</td>
<td>Live birth</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design</td>
<td>N Enrolled</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>------------</td>
<td>---------------</td>
<td>------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ 3</td>
<td>RCT</td>
<td>180 170 Unknown</td>
<td>5 mg/day letrozole on day 3-7 of menstrual cycle vs. 100 mg/day clomiphene on day 3-7 of menstrual cycle</td>
<td>Miscarriage OHSS Cost</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>In both groups, human menopausal gonadotropin was administered every day starting on day between 6-8 of cycle. Ovulation was triggered with urinary Human Chorionic Gonadotropin (5000 IU) when have two follicles of ≥16 mm. IUI was performed 36 hr later</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Across All KQs</td>
<td>Observational US</td>
<td>Treatment received in US states that mandated coverage for IVF vs. Treatment received in US states that did not mandate coverage for IVF</td>
<td>Live birth Multiple births Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asia</td>
<td>481 442 Tubal</td>
<td>IVF + TEAS-2Hz vs. IVF + TEAS-100 Hz vs. IVF + TEAS-2/100Hz vs. IVF</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>UK/Europe</td>
<td>304 249 Unknown</td>
<td>CC oral tablets at dose of 150 mg/d from day 3 to day 7 vs. Daily SQ injections of triptoreline started on day one or two of the menstrual cycle and 450 IU of SQ rFSH from day 3 of the cycle.</td>
<td>Live birth (ovarian reserve) Patient costs</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQs Companion Studies</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Studies Design Geographical Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rahman, 2017</td>
<td>RCT</td>
<td>UK/Europe</td>
<td>66 61 Male</td>
<td>Recombinant FSH (r-FSH) supplemented by r-LH in the late follicular phase starting the same day of GnRH-antagonist (GnRH-ant) administration vs. r-FSH alone</td>
<td>Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rashidi, 2015</td>
<td>RCT</td>
<td>Middle East</td>
<td>276 276 PCOS</td>
<td>Ovulation induction with CC day 3-7 followed by: 75 IU of rFSH day 7-9 and IUI vs. 75 IU of hMG day 7-9 and IUI</td>
<td>Live birth Miscarriage Multiple births OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rashidi, 2013</td>
<td>RCT</td>
<td>Middle East</td>
<td>280 259 Unknown</td>
<td>Induction of ovulation with CC 100 mg/d on cycle days 3 –7, followed by 75 IU/d hMG on cycle days 7–9, followed by triggered ovulation and IUI vs. Same protocol except hMG replaced by 75 IU rFSH.</td>
<td>Live birth Miscarriage OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Razi, 2013</td>
<td>RCT</td>
<td>Middle East</td>
<td>182 182 Male</td>
<td>Laser Assisted Hatching (LAH) following ICSI vs. Control group. Intact transferred embryos without LAH</td>
<td>Live birth Multiple births Congenital anomalies</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reindollar, 2010</td>
<td>RCT</td>
<td>US</td>
<td>503 503 Unknown</td>
<td>CC/IUI x 3 cycles followed by up to 3 cycles gonadotropin/IUI followed by 6 cycles of IVF (of which 2 could be frozen cycles) vs. CC/IUI x 3 cycles followed by 6 cycles of IVF (of which 2 could be frozen cycles)</td>
<td>Live birth Ectopic pregnancy Miscarriage Multiple births Time to pregnancy OHSS Birthweight Neonatal death Health system costs</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>----------------------------------</td>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rubio, 2013&lt;sup&gt;217&lt;/sup&gt; KQ 3</td>
<td>RCT UK/Europe</td>
<td>274</td>
<td>IVF with day-5 blastocyst transfer (no PGD) vs. IVF, embryo biopsy and FISH for 9 chromosomes on day 3, transfer on day 5</td>
<td>Live birth Miscarriage Multiple births (RIF and AMA)</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schendelaar, 2011&lt;sup&gt;273&lt;/sup&gt; Across All KQs</td>
<td>Observational UK/Europe</td>
<td>NR 310</td>
<td>Children who were not born to sub-fertile parents vs. Children born to sub-fertile parents who received IVF (either COH-IVF or MNC-IVF)</td>
<td>Neurodevelopmental issues</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seckin, 2014&lt;sup&gt;209&lt;/sup&gt; KQ 3</td>
<td>RCT UK/Europe</td>
<td>149</td>
<td>Luteal phase support with 90 mg/d vaginal 8% progesterone gel starting on the day of IUI until pregnancy testing vs. No drug for luteal phase support</td>
<td>Live birth</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selman, 2016&lt;sup&gt;237&lt;/sup&gt; KQ 3</td>
<td>RCT UK/Europe</td>
<td>85</td>
<td>Clomiphene citrate and corifollitropin alfa for the first 7 days of stimulation followed by recombinant follicle stimulating hormone (rFSH) in a gonadotropin-releasing hormone antagonist protocol vs. Clomiphene citrate and a daily injection of rFSH in a gonadotropin-releasing hormone antagonist protocol</td>
<td>Live birth Miscarriage</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shi, 2018&lt;sup&gt;262&lt;/sup&gt; KQ 4 and 5</td>
<td>RCT Asia</td>
<td>2,157</td>
<td>Fresh-embryo transfer vs. Embryo cryopreservation followed by frozen-embryo transfer</td>
<td>Live birth Multiple births Miscarriage OHSS Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sismanoglu, 2009&lt;sup&gt;267&lt;/sup&gt; KQ 6</td>
<td>RCT UK/Europe</td>
<td>50</td>
<td>Donor triggering with hCG vs. Donor triggering with GnRH agonist</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>---------------------------------</td>
<td>-------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spaan, 2015\textsuperscript{125}</td>
<td>Observational UK/Europe</td>
<td>25,108 25,108 All</td>
<td>Received at least one IVF cycle with ovarian stimulation vs. Subfertile women with other treatments including tubal surgery, IUI, or hormonal treatment OR withdrew from IVF waiting list</td>
<td>Maternal cancer</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OMEGA Across All KQs Companions: Spaan, 2016\textsuperscript{127}; van Leeuwen, 2011\textsuperscript{126}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stadtmauer, 2011\textsuperscript{159} KQ 1</td>
<td>RCT US</td>
<td>98 NR PCOS</td>
<td>rFSH Follistim starting on cycle day 3 until the appropriate follicle size was reached. vs. Ganirelix 0.25 mg SQ/d added to rFSH in a flexible protocol when the leading follicle diameter reached R13 mm. vs. rFSH and Ganirelix</td>
<td>Live birth Miscarriage Multiple births</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stewart, 2013\textsuperscript{203} KQ 2</td>
<td>Observational Australia/NZ</td>
<td>22,045 21,646 Endometriosis</td>
<td>Women receiving infertility treatment but not IVF vs. Women receiving IVF</td>
<td>Maternal cancer (parity)</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tartagni, 2015\textsuperscript{207} KQ 3</td>
<td>RCT UK/Europe</td>
<td>109 109 Unknown</td>
<td>IVF with DHEA 75 mg/d for 8 weeks pre-treatment vs. Placebo</td>
<td>Live birth Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tehrani nejad, 2010\textsuperscript{151} KQ 1</td>
<td>RCT Middle East</td>
<td>95 90 PCOS</td>
<td>Pretreatment with OCP for 21 days, controlled ovarian stim started on day 2-3. Cetrotix (antagonist) 0.25 mg SQ started when follicles 12-14 mm. vs. Control group. Pretreatment with OCP for 21 days, along with buserelin (agonist) 500 mcg/d SQ. Buserelin then reduced to 250 mcg/d. Controlled ovarian stimulation with hMG.</td>
<td>OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------</td>
<td>------------------------------------------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toftager, 2016&lt;sup&gt;7,3&lt;/sup&gt;</td>
<td>RCT UK/Europe</td>
<td>1,099 1,023 All</td>
<td>GnRH antagonist protocol: Women with regular cycles (≤35 days) received daily injections with recombinant human FSH. Controlled ovarian stimulation (COS) was initiated on cycle day 2 or 3 and continued until the day of ovulation induction. A fixed dose of rFSH was used for the first 6 days, either 150 or 225 IU according to age ≤36 years or age &gt;36 years. After 6 days of stimulation the rFSH doses were adjusted according to ovarian response evaluated using transvaginal ultrasonography. A daily GnRH antagonist (Ganirelix) dose of 0.25 mg was used starting on stimulation day 6 and was administered until the day of ovulation induction. vs.</td>
<td>Live birth OHSS</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toftager, 2017&lt;sup&gt;7,3&lt;/sup&gt;</td>
<td>RCT UK/Europe</td>
<td>1,050 1,023 All</td>
<td>Short GnRH-antagonist protocol vs. Long GnRH-agonist protocol</td>
<td>Live birth Miscarriage Time to pregnancy</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>KQs</td>
<td>Study Design</td>
<td>Geographic Location</td>
<td>N Enrolled</td>
<td>N Completed</td>
<td>N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
</tr>
<tr>
<td>---------------</td>
<td>-----</td>
<td>--------------</td>
<td>---------------------</td>
<td>------------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>--------------</td>
<td>--------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Topcu, 2017</td>
<td>KQ 1</td>
<td>RCT</td>
<td>Middle East</td>
<td>101</td>
<td>88</td>
<td>PCOS</td>
<td>Tamoxifen 20mg/daily for 5 consecutive days vs. Clomiphene citrate 50mg/daily for 5 consecutive days Both medications were started on day 5 of the menstrual cycle</td>
<td>Live birth Miscarriage Multiple births OHSS</td>
<td>Fair</td>
</tr>
<tr>
<td>Tsai, 2011</td>
<td>KQ 5</td>
<td>Observational</td>
<td>Asia</td>
<td>NA</td>
<td>191 cycles Male</td>
<td>Testicular sperm extraction (TESE) via biopsy in men w/ azoosperma vs. Fresh ejaculated sperm from men with extreme severe oligo-asthenoteratozoospermia (OAT) sperm</td>
<td>Live birth Miscarriage Birthweight Congenital anomalies</td>
<td>Poor</td>
<td></td>
</tr>
<tr>
<td>van Rijswijk, 2017</td>
<td>KQ 3</td>
<td>RCT</td>
<td>UK/Europe</td>
<td>498</td>
<td>481 Unknown</td>
<td>Live birth Multiple births Miscarriage Time to pregnancy</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Rumste, 2014</td>
<td>KQ 3</td>
<td>RCT</td>
<td>UK/Europe</td>
<td>116</td>
<td>116 Unknown</td>
<td>IVF elective single embryo transfer vs. COH/IUI, maximum treatment duration 3 cycles</td>
<td>Live birth Health system costs</td>
<td>Fair</td>
<td></td>
</tr>
<tr>
<td>Verhoeve, 2013</td>
<td>KQ 4</td>
<td>Observational</td>
<td>UK/Europe</td>
<td>5000</td>
<td>5000 Tubal Factor</td>
<td>No diagnostics and no treatment vs. No diagnostics and immediate treatment (up to 3 IVF treatments) vs. No diagnostics and delayed treatment (no treatment for 1 yr, then up to 3 IVF treatments) vs. Hysterosalpingogram followed by tailored treatment (delayed or immediate IVF)</td>
<td>Live birth Patient costs</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------------</td>
<td>-----------------------------</td>
<td>--------------</td>
<td>-----------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, 2016</td>
<td>RCT Asia</td>
<td>120 108 PCOS</td>
<td>hMG 150-225IU/day and MPA 10mg/day from day 3 of menstruation. Both cotriggered with GnRHa and hCG and underwent IVF/ICSI. vs. Patients received GnRHa (decapeptyl) 0.1mg beginning of day 2 of menstruation and adding hMG on day 3 of menstruation. Both cotriggered with GnRHa and hCG and underwent IVF/ICSI.</td>
<td>Multiple births Miscarriage Ectopic pregnancy OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wang, 2017</td>
<td>Observational US</td>
<td>509,938 cycles 509,938 cycles All</td>
<td>Fresh blastocyst vs. Fresh non-blastocyst vs. Frozen blastocyst vs. Frozen non-blastocyst</td>
<td>Live birth Miscarriage Ectopic pregnancy</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiss, 2018</td>
<td>RCT UK/Europe</td>
<td>666 661 PCOS, Unknown</td>
<td>Six cycles with gonadotrophins plus intrauterine insemination vs. Six cycles with gonadotrophins plus intercourse vs. Six cycles with clomifene citrate plus intrauterine insemination vs. Six cycles with clomifene citrate plus intercourse Clomifene citrate dosages varied from 50 to 150 mg daily orally and gonadotrophin starting dose was 50 or 75 IU daily subcutaneously</td>
<td>Live birth Multiple births Miscarriage Ectopic pregnancy Time to pregnancy Birthweight</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Williams, 2018</td>
<td>Observational UK/Europe</td>
<td>255,786 255,786 All</td>
<td>Assisted Reproduction vs. General Population</td>
<td>Maternal cancer</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design</td>
<td>Geographic Location</td>
<td>N Enrolled</td>
<td>N Completed</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>---------------------</td>
<td>------------</td>
<td>-------------</td>
<td>---------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wiser, 2010²²⁰</td>
<td>RCT</td>
<td>Middle East</td>
<td>33</td>
<td>33</td>
<td>Long-stimulation protocol IVF with 75 mg DHEA orally, once a day, at least 6 weeks before starting the first cycle of ovulation induction. Patients who did not conceive and continued to the second cycle took DHEA for at least 16–18 weeks. vs. Standard long-stimulation protocol IVF</td>
<td>Live birth (ovarian reserve)</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu, 2016¹⁷⁵</td>
<td>RCT</td>
<td>Asia</td>
<td>644</td>
<td>644 in analysis</td>
<td>Letrozole 2.5 mg plus placebo for Berberine, increased to 5.0 mg on days 3-7 of last 3 treatment cycles vs. Berberine 1.5 mg plus placebo for Letrozole vs. Combination letrozole and berberine</td>
<td>Live birth Multiple births Miscarriage Ectopic pregnancy Birthweight Birthweight Neonatal death</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wu, 2017¹⁸⁸</td>
<td>RCT</td>
<td>Asia</td>
<td>1,000</td>
<td>926</td>
<td>Active acupuncture administered twice a week plus clomiphene administered for 5 days per cycle, for up to 4 cycles vs. Active acupuncture administered twice a week plus placebo for clomiphene administered for 5 days per cycle, for up to 4 cycles vs. Control acupuncture administered twice a week plus clomiphene administered for 5 days per cycle, for up to 4 cycles vs. Control acupuncture administered twice a week plus placebo for clomiphene administered for 5 days per cycle, for up to 4 cycles</td>
<td>Live birth Multiple births Miscarriage Ectopic pregnancy Birthweight Neonatal death</td>
<td>Good</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xiong, 2017²⁶³</td>
<td>Observational</td>
<td>US</td>
<td>141,030</td>
<td>141,030 Male</td>
<td>IVF vs. ICSI</td>
<td>Neonatal death Preterm birth Congenital anomaly</td>
<td>Fair</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yapca, 2015\textsuperscript{206} KQ 3</td>
<td>RCT UK/Europe</td>
<td>80 80 Unknown</td>
<td>CC 100 mg/d on cycle days 3 - 7 followed by time-limited hydrotubation performed after detection of the dominant follicle and then timed intercourse vs. No hydrotubation</td>
<td>Live birth Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yazici, 2014\textsuperscript{144} KQ 1</td>
<td>RCT UK/Europe</td>
<td>110 56 PCOS</td>
<td>Ovarian stimulation using rFSH followed by IUI and luteal support with vaginal micronized progesterone 300 mg/d vs. No luteal support</td>
<td>Live birth</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yildiz, 2014\textsuperscript{211} KQ 3</td>
<td>RCT UK/Europe</td>
<td>180 180 Unknown</td>
<td>Follitropin alpha, follitropin beta, uhMG and urofolitropin were used for ovarian stimulation. Ovulation induction was started between 2-5 days of menstruation on patients who had no residual cysts larger than 15 mm as visualized with basal transvaginal USG (ultrasound). All patients had 75-150 IU/d drug as an initial dose. On cycle day 5-6, stimulated follicles were measured ultrasonographically. Induction doses were increased or decreased between 37.5-75 IU/d according to follicle size. When 1-2 follicles reached a mean diameter of 17 mm, 250 mcg of rhCG was administered to trigger ovulation. Uterine washing was accomplished by introducing a silicone catheter through the internal cervical os, after which 20 cc saline and 1 cc jetocain were slowly injected. The speculum was removed and the procedure completed after the injection. At 35-36 hours after the hCG injection, IUI was performed. vs. Same procedures except no uterine washing performed</td>
<td>Live birth Miscarriage</td>
<td>Poor</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Youssef, 2016</td>
<td>RCT Africa</td>
<td>394 394 Unknown</td>
<td>Mild ovarian stimulation. Pretreatment with an oral contraceptive pill was followed by ovarian stimulation starting with a fixed daily dose of 150 IU/day FSH on Day 5. On stimulation Day 6, 0.25 mg/day s.c. of a GnRH antagonist (Cetrotide®) was commenced. Ovulation was triggered by 10,000 IU human chorionic gonadotropin hormone (Pregnyl) when a leading follicle reached 18 mm, and follicle aspiration was done by transvaginal ultrasound guided oocyte retrieval 34–36 h thereafter. Subsequently, embryo transfers were performed according to the local policy vs. Conventional stimulation protocol. In the women allocated to the conventional ovarian stimulation strategy, daily injections were given of 0.1 mg s.c of a gonadotropin releasing hormone agonist to prevent premature ovulation (Decapeptyl®) followed by stimulation with fixed daily injections of 450 IU HMG (Menopur® or Merional®). Ovulation was triggered by 10,000 IU human chorionic gonadotropins hormone (Pregnyl) when a leading follicle reached 18 mm and follicle aspiration was done by transvaginal ultrasound guided oocyte retrieval 34–36 h thereafter. The remainder of the cycle was identical to the mild ovarian stimulation strategy</td>
<td>Miscarriage</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yu, 2018</td>
<td>RCT Asia</td>
<td>116 116 Unknown</td>
<td>Modified GnRH agonist (triptorelin) vs. Mild stimulation protocol with letrozole vs. Antagonist protocol with triptorelin</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zahran, 2018</td>
<td>RCT Middle East</td>
<td>130 120 PCOS</td>
<td>CC + Cabergoline vs. CC alone</td>
<td>Miscarriage OHSS</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM</td>
<td>Study Design</td>
<td>N Enrolled</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>------------</td>
<td>--------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zain, 2009[^166]</td>
<td>RCT Asia</td>
<td>124 PCOS</td>
<td>Metformin tablets at the initial dose of 500 mg and increased in a step-wise fashion during the first 3 weeks to a total dose of 1,500 mg/d. The patients were then asked to make a telephone call once they had a menstrual period and a transvaginal ultrasound (TVS) and follicular tracking was done to document evidence of follicular growth and ovulation on days 2, 8, 12, and 16. A menstrual calendar chart recorded menses cycles monthly. vs. CC at a dose of 50 mg on days 2–6. The TVS and follicular tracking were done to document follicular growth and ovulation on days 2, 8, 12, and 16. If there was absence of ovulation, the CC dose was increased stepwise on a treatment cycle basis after a P withdrawal bleed to a maximum of 200 mg. If there was evidence of ovulation but the patient did not get pregnant, the same dosage was continued for a maximum of six cycles. vs. Metformin was given in a similar manner to the metformin only group. CC was given at a dose of 50 mg on days 2–6. The TVS and follicular tracking were done to document evidence of follicular growth and ovulation on days 2, 8, 12, and 16. If there was absence of ovulation, the CC dose was increased step-wise on a treatment cycle basis after a P withdrawal bleed to a maximum of 200 mg. If there was evidence of ovulation but patient did not get pregnant, a similar dosage was continued for a maximum of six cycles.</td>
<td>Live birth Ectopic pregnancy Miscarriage Multiple births</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KQ 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study ACRONYM KQs Companion Studies</td>
<td>Study Design Geographic Location</td>
<td>N Enrolled N Completed Underlying Diagnosis</td>
<td>Interventions</td>
<td>Outcomes (Subgroups analyzed)</td>
<td>Quality</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>---------------------------------</td>
<td>--------------------------------------------</td>
<td>---------------</td>
<td>-------------------------------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zakherah, 2010&lt;sup&gt;169&lt;/sup&gt;</td>
<td>RCT Africa</td>
<td>150&lt;br&gt;150&lt;br&gt;PCOS</td>
<td>CC 150 mg + Tamoxifen 40 mg from cycle days 3 to 7, maximum treatment duration 6 cycles. vs. Laparoscopic ovarian drilling (LOD) performed through triple-puncture laparoscopy followed by timed intercourse.</td>
<td>Live birth Miscarriage</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zhu, 2014&lt;sup&gt;202&lt;/sup&gt;</td>
<td>RCT Asia</td>
<td>163&lt;br&gt;156&lt;br&gt;Endometriosis</td>
<td>Oral contraceptive for 63 days, every 3 month visits for 14 months, if no pregnancy within 12 months of stopping OCP, advised to undergo IVF vs. Oral contraceptive for 33 days, followed by a combination of oral contraceptive and 30 g/d Dan’e mixture for 30 days; every 3 month visits for 14 months, if no pregnancy within 12 months of stopping OCP, advised to undergo IVF vs. No treatment; q3month visits for 12 months, if no pregnancy within 12 months, advised to undergo IVF</td>
<td>Live birth Ectopic pregnancy Miscarriage Neonatal death</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zarei, 2016&lt;sup&gt;211&lt;/sup&gt;</td>
<td>RCT Middle East</td>
<td>260&lt;br&gt;260&lt;br&gt;Unknown</td>
<td>Clomid days 5-9, recombinant FSH days 8+, hCG trigger, IUI x up to 3 cycles. Piroxicam days 4-6 after IUI. vs. Clomid days 5-9, recombinant FSH days 8+, hCG trigger, IUI x up to 3 cycles.</td>
<td>Miscarriage</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zheng, 2012&lt;sup&gt;149&lt;/sup&gt;</td>
<td>RCT Asia</td>
<td>82&lt;br&gt;74&lt;br&gt;PCOS</td>
<td>Primed with 10,000 IU hCG after progesterone induced withdrawal bleeding. Immature oocytes were collected 36-38 hours after hCG priming. IVM and ICSI were done. vs. No priming after progesterone induced withdrawal bleeding. Immature oocytes were collected directly after allocation to non-priming group. IVM and ICSI were done.</td>
<td>Live birth</td>
<td>Fair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Abbreviations: ART=assisted reproductive technology; BID=two times per day; BMI=body mass index; CARE Consortium=Centres for Assisted Reproduction; CC=clomiphene citrate; COH=controlled ovarian hyperstimulation; COS=controlled ovarian stimulation; DHEA=dehydroepiandrosterone; FASTT=Fast Track and Standard Treatment Trial; FISH=fluorescence in situ hybridization; FORT-T=Forty and Over Treatment Trial; FSH=follie stimulating hormone; GnRH=growth hormone-releasing hormone; hCG=human chorionic gonadotropin; hMG=human menopausal gonadotropin; ICSI=intracytoplasmic sperm injection; IMSI=intracytoplasmic morphologically selected sperm injection; ITT=intention-to-treat; IU=international units; IUI=intrauterine insemination; IVF=in vitro fertilization; IVM=in vitro maturation; KQ=key question; mcg=microgram; MNC=modified natural cycle; MOSART=Massachusetts Outcomes Study of Assisted Reproductive Technologies; NA=not applicable; NASS=National Artificial Reproductive Technology Surveillance System; NR=Not Reported; OCP=oral contraceptive pill; OHSS=Ovarian Hyperstimulation Syndrome; PCOS=Polycystic Ovary Syndrome; PGD=preimplantation genetic diagnosis; PPCOS=Pregnancy in Polycystic Ovary Syndrome; RCT=Randomized Controlled Trial; rFSH=recombinant follicle stimulating hormone; rhFSH=recombinant human follicle stimulating hormone; SQ=subcutaneous; SUIT=Scottish Unexplained Infertility Trial; TVS=transvaginal ultrasound; uFSH=urinary follicle stimulating hormone; uhMG=urinary human menopausal gonadotropin
Appendix F. AMSTAR Quality Assessment for Systematic Reviews

Table F-1 shows the AMSTAR (A Measurement Tool to Assess the Methodological Quality of Systematic Reviews) quality assessment for the included systematic reviews. For full study citations, please refer to the report’s main reference list.

<table>
<thead>
<tr>
<th>Study</th>
<th>&quot;A Prior&quot; design provided?</th>
<th>Duplicate study selection and data abstraction?</th>
<th>Comprehensive literature search performed?</th>
<th>Status of publication (i.e. grey literature used as an inclusion criterion)?</th>
<th>List of studies (included and excluded) provided?</th>
<th>Characteristics of included studies provided?</th>
<th>Scientific quality of included studies assessed and documented?</th>
<th>Scientific quality of the included studies used appropriately in formulating conclusions?</th>
<th>Methods used to combine the findings of studies appropriate?</th>
<th>Likelihood of publication bias assessed?</th>
<th>Conflict of interest included?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Al-Inany, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Attia, 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Barbosa, 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Bensdorp, 2007</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Bhattacharya, 2010</td>
<td>Y</td>
<td>C</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Bordewijk, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Brown, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Brown, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Brown, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Chua, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Cissen, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Duffy, 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Farquhar, 2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Farquhar, 2012</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Farquhar, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Franik, 2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Franik, 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Franik, 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Ghobara, 2017</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hamdan, 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hamdan, 2015</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hu, 2018</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Hughes, 2010</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Study</td>
<td>&quot;A Priori&quot; design provided?</td>
<td>Duplicate study selection and data abstraction?</td>
<td>Comprehensive literature search performed?</td>
<td>Status of publication (i.e. grey literature) used as an inclusion criterion?</td>
<td>List of studies (included and excluded) provided?</td>
<td>Characteristics of included studies provided?</td>
<td>Scientific quality of included studies assessed and documented?</td>
<td>Scientific quality of the included studies used appropriately in formulating conclusions?</td>
<td>Methods used to combine the findings of studies appropriate?</td>
<td>Likelihood of publication bias assessed?</td>
<td>Conflict of interest included?</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
<td>------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Jie, 2018[31]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Johnson, 2010[93]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Lensen, 2016[12]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Lim, 2016[105]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Liu, 2014[70]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Morley, 2017[199]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Mourad, 2017[133]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Nagels, 2015[134]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Noble, 2010[94]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>C</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Osman, 2015[61]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Pandian, 2012[84]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Pandian, 2015[515]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Pandian, 2010[75]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Reavey, 2016[106]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Showell, 2011[136]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Siristatidis, 2016[157]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Siristatidis, 2015[158]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sun, 2013[78]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>C</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Tang, 2016[139]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Teixeira, 2013[39]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Veltman-Verhulst, 2012[87]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Veltman-Verhulst, 2016[140]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Vermy, 2018[141]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Vitaglia, 2018[142]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Youssef, 2016[143]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Youssef, 2015[144]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Yu, 2018[145]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Zhou, 2016[146]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Abbreviations: C=Can’t answer; N=No; Y=Yes
Figure F-1. Summary of AMSTAR quality assessment for systematic reviews

- "A Priori" design provided?
- Duplicate study selection and data abstraction?
- Comprehensive literature search performed?
- Status of publication (i.e. grey literature) used as an inclusion criterion?
- List of studies (included and excluded) provided?
- Characteristics of included studies provided?
- Scientific quality of included studies assessed and documented?
- Scientific quality of the included studies used appropriately in formulating conclusions?
- Methods used to combine the findings of studies appropriate?
- Likelihood of publication bias assessed?
- Conflict of interest included?

Percent of studies with low, high, or unclear risk of bias

- Yes
- No
- Can't Answer
Appendix G. Risk of Bias Assessment for Included Studies

Table G-1 shows the risk of bias quality assessment for the included cohort studies. For full study citations, please refer to the report’s main reference list.

Table G-1. Risk of bias assessment for included cohort studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Were participants analyzed within the groups they were originally assigned to?</th>
<th>Did the study apply inclusion criteria uniformly to all comparison groups?</th>
<th>Did the study maintain fidelity to the intervention protocol?</th>
<th>If attrition was a concern, was missing data handled appropriately?</th>
<th>Were the outcome assessors blinded to the status of participants?</th>
<th>Was the length of follow-up the same between the groups?</th>
<th>Were intervention/exposures implemented consistently across all study participants?</th>
<th>Were outcomes assessed, implemented consistently across all study participants?</th>
<th>Were confounding variables adjusted for?</th>
<th>Were potential outcomes reported by the researchers?</th>
<th>Are all reported outcomes prespecified?</th>
<th>Reporting bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belva, 2011–17</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Bodri, 2008–12</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Bodri, 2009–13</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Boulet, 2015–19</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Boulet, 2016–20</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Brinton, 2015–19</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Butts, 2014–18</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Chang, 2016–20</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Crawford, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>de Wilde, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Dhalwani, 2016–20</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Hershko-Klement, 2016–20</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>Kettner, 2016–20</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Keyhan, 2018–22</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Kissin, 2015–19</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Knudtson, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Levi Dunietz, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Litzy, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Litzy, 2018–22</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Luka, 2010–22</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Magnusson, 2016–22</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Malchau, 2017–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Mancuso, 2016–21</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Maxwell, 2008–22</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>Muller, 2017–22</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>

G-1
| Study                | Were participants analyzed within the groups they were originally assigned to? | Did the study apply inclusion/exclusion criteria uniformly to all comparison groups? (selection bias) | Did the strategy for recruiting participants differ across study groups? (selection bias) | Did the design or analysis account for confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches? (selection bias) | Did researchers rule out any intervention or unintended exposure that might bias results? (performance bias) | Did the study maintain fidelity to the intervention protocol? (performance bias) | If attrition was a concern, were missing data handled appropriately? (performance bias) | Were the length of follow-up the same between the groups? (detection bias) | Were the outcome assessors blinded to the status of participants? (detection bias) | Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias) | Were potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias) |
|---------------------|--------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|---------------------------------------------------------------------------------|
| Nangia, 2011\(^{24}\) | U                                                                              | Y                                                                                               | Y                                                                                   | N                                                                              | Y                                                                              | N                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Oyesanya, 2009\(^{24}\) | Y                                                                              | Y                                                                                               | N                                                                                   | N                                                                              | Y                                                                              | Y                                                                              | N                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              |
| Schendelaar, 2011\(^{23}\) | Y                                                                              | Y                                                                                               | Y                                                                                   | U                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | U                                                                              |
| Spaan, 2015\(^{16}\) | N                                                                              | Y                                                                                               | N                                                                                   | N                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              | Y                                                                              | Y                                                                              | U                                                                              |
| Stewart, 2013\(^{26}\) | Y                                                                              | Y                                                                                               | N                                                                                   | N                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              |
| Tsai, 2011\(^{27}\) | N                                                                              | N                                                                                               | Y                                                                                   | N                                                                              | N                                                                              | U                                                                              | N                                                                              | N                                                                              | N                                                                              | Y                                                                              | N                                                                              |
| Verhoeve, 2013\(^{27}\) | U                                                                              | U                                                                                               | U                                                                                   | U                                                                              | U                                                                              | U                                                                              | U                                                                              | U                                                                              | U                                                                              | U                                                                              | U                                                                              |
| Vitek, 2013\(^{23}\) | Y                                                                              | Y                                                                                               | N                                                                                   | N                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              |
| Wang, 2017\(^{24}\) | Y                                                                              | Y                                                                                               | N                                                                                   | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | U                                                                              | Y                                                                              |
| Williams, 2013\(^{21}\) | Y                                                                              | Y                                                                                               | N                                                                                   | N                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | U                                                                              | Y                                                                              |
| Williams, 2018\(^{20}\) | Y                                                                              | Y                                                                                               | N                                                                                   | U                                                                              | Y                                                                              | N                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Xiong, 2017\(^{20\dagger}\) | Y                                                                              | Y                                                                                               | U                                                                                   | Y                                                                              | N                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |

Abbreviations: N=No; U=Unclear; Y=Yes
Figure G-1. Summary of risk of bias assessment for included cohort studies

Were participants analyzed within the groups they were originally assigned to? (selection bias)

Did the study apply inclusion/exclusion criteria uniformly to all comparison groups? (selection bias)

Did the strategy for recruiting participants into the study differ across study groups? (selection bias)

Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or…

Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? (performance bias)

Did the study maintain fidelity to the intervention protocol? (performance bias)

If attrition was a concern, were missing data handled appropriately? (attrition bias)

In prospective studies, was the length of follow-up the same between the groups, or in case-control studies, was the time period between the…

Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias)

Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)

Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)

Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants? (detection bias)

Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias)
Table G-2 shows the risk of bias quality assessment for the included cross-sectional studies. For full study citations, please refer to the report’s main reference list.

**Table G-2. Risk of bias assessment for included cross-sectional studies**

<table>
<thead>
<tr>
<th>Study</th>
<th>Did the study apply inclusion/exclusion criteria uniformly to all comparison groups? (selection bias)</th>
<th>Does the design or analysis control and adjust for important confounding variables through matching, stratification, or multivariable analysis? (selection bias)</th>
<th>Did researchers rule out any impact from a concurrent intervention or unintended exposure that might bias results? (performance bias)</th>
<th>If attrition was a concern, were missing data handled appropriately? (attrition bias)</th>
<th>Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias)</th>
<th>Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)</th>
<th>Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)</th>
<th>Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants? (detection bias)</th>
<th>Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barad, 2017</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Grimstad, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Kramer, 2009</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Londra, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Luke, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Provost, 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Abbreviations: N=No; U=Unclear; Y=Yes
Figure G-2. Summary of risk of bias assessment for included cross-sectional studies

- Did the study apply inclusion/exclusion criteria uniformly to all comparison groups? (selection bias)
- Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches? (selection bias)
- Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? (performance bias)
- If attrition was a concern, were missing data handled appropriately? (attrition bias)
- Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias)
- Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)
- Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)
- Were confounding variables assessed using valid and reliable measures, implemented consistently across all study participants? (detection bias)
- Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias)
Table G-3 shows the risk of bias quality assessment for the included RCTs. For full study citations, please refer to the report’s main reference list.

| Study                                      | Was the allocation sequence generated adequately? (selection bias) | Was the allocation of treatment adequately concealed? (selection bias) | Were participants analyzed within the groups they were originally assigned to? | Did the design or analysis account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches? (selection bias) | Did researchers rule out any impact from a conflict of interest or any other potential bias? | Did the study maintain fidelity to the intervention protocol (performance bias) | If attrition was a concern, were missing data handled appropriately? (attrition bias) | In prospective studies, was the length of follow-up the same between groups, or was intervention or exposure status confounded by protocol violations? | Were the outcome assessors blinded to the intervention or exposure status of all study participants? (detection bias) | Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias) |
|--------------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|--------------------------------------------------------------------------------|
| Abdellah, 2011[16]                         | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Aboulghar, 2010[6]                         | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Abu Hashim, 2012[17]                       | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Abu Hashim, 2011[15]                       | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Abu Hashim, 2010[14]                       | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Abu Hashim, 2017[13]                       | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Aghahosseini, 2017[18]                     | Y                                                            | U                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Ameer, 2017[19]                            | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Amer, 2009[16]                             | Y                                                            | Y                                                            | N                                                                              | Y                                                                              | N                                                                              | N                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y N                                                                              |
| An, 2014[17]                               | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Badawy, 2008[17]                           | Y                                                            | U                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Badawy, 2009[22]                           | N                                                            | N                                                            | N                                                                              | N                                                                              | Y                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Bagis, 2010[21]                            | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Balaban, 2011[23]                          | U                                                            | U                                                            | Y                                                                              | N                                                                              | U                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Bhattacharya, 2008[13]                     | Y                                                            | U                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Chen, 2016[18]                             | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Choi, 2012[14]                             | U                                                            | U                                                            | Y                                                                              | U                                                                              | U                                                                              | U                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Custers, 2012[19]                          | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | U                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Danhof, 2018[14]                           | N                                                            | N                                                            | Y                                                                              | N                                                                              | U                                                                              | U                                                                              | U                                                                              | Y                                                                              | Y                                                                              | U                                                                              |
| Demiroi, 2007[13]                          | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Diamond, 2015[15]                          | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Ebrahimi, 2010[25]                         | Y                                                            | Y                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |
| Einarsson, 2017[14]                        | Y                                                            | U                                                            | Y                                                                              | Y                                                                              | Y                                                                              | Y                                                                              | N                                                                              | Y                                                                              | Y                                                                              | Y                                                                              |

G-6
<p>| Study                           | Was the allocation sequence generated adequately? | Was the allocation of treatment adequately concealed? | Were participants analyzed within the groups they were originally assigned to? | Did researchers rule out any impact from a confounding or modifying variable through matching, stratifying, or matching another variable? | Did the study maintain fidelity to the intervention protocol? (performance bias) | Did attrition concern, if any, missing data handled appropriately? | In prospective studies, was the length of follow-up the same between the groups or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls? | Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias) | Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias) | Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (reporting bias) | Were the potential outcomes specified by the researchers? Are all prespecified outcomes reported? (reporting bias) |
|--------------------------------|-------------------------------------------------|-----------------------------------------------------|---------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Elsedeek, 2014[43]            | Y                                               | Y                                                   | U                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Emekci Ozay, 2017[42]         | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Erdem, 2009[45]               | Y                                               | N                                                   | Y                                                                               | N                                                                               | U                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | N                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Erdem, 2015[47]               | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Farquhar, 2018[48]            | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                                  | N                                                                                  |
| Ge, 2008[50]                  | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | U                                                                               | Y                                                                               | U                                                                                  | U                                                                                  | Y                                                                                  | Y                                                                                  |
| Ghahiri, 2016[51]             | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | N                                                                                  | U                                                                               | U                                                                               | Y                                                                                  | U                                                                                  | Y                                                                                  | N                                                                                  |
| Ghanem, 2013[49]              | Y                                               | N                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | U                                                                               | N                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Gibreel, 2013[52]             | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | N                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Goldman, 2014[14]             | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Gregoriou, 2008[53]           | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Hajizadeh, 2017[54]           | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | N                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Hantra, 2018[40]              | Y                                               | Y                                                   | Y                                                                               | U                                                                               | U                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | U                                                                                  | Y                                                                                  |
| Hassan, 2017[55]              | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Homburg, 2012[56]             | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Hossein, 2010[57]             | U                                               | U                                                   | U                                                                               | N                                                                               | U                                                                                  | U                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Hossein-Rashidi, 2016[58]     | Y                                               | U                                                   | U                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | N                                                                               | U                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Ibrahim, 2017[59]             | Y                                               | U                                                   | Y                                                                               | N                                                                               | Y                                                                                  | U                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | U                                                                                  |
| Jacob, 2016[60]               | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Jahromi, 2017[61]             | Y                                               | Y                                                   | Y                                                                               | N                                                                               | U                                                                                  | U                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Johnson, 2010[62]             | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kansal Kaira, 2008[63]        | Y                                               | Y                                                   | Y                                                                               | N                                                                               | U                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kar, 2015[64]                 | N                                               | Y                                                   | Y                                                                               | N                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | N                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Khosravi, 2015[65]            | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | U                                                                                  | U                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kim, 2011[66]                 | U                                               | U                                                   | Y                                                                               | U                                                                               | U                                                                                  | U                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kim, 2012[67]                 | Y                                               | Y                                                   | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kjotred, 2011[68]             | Y                                               | Y                                                   | Y                                                                               | N                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |
| Kurzawa, 2008[69]             | Y                                               | N                                                   | Y                                                                               | Y                                                                               | U                                                                                  | Y                                                                                  | Y                                                                               | N                                                                               | Y                                                                                  | U                                                                                  | Y                                                                                  | Y                                                                                  |
| Kuzmin, 2014[70]              | U                                               | U                                                   | U                                                                               | N                                                                               | U                                                                                  | N                                                                                  | U                                                                               | Y                                                                               | Y                                                                                  | Y                                                                                  | Y                                                                                  | Y                                                                                  |</p>
<table>
<thead>
<tr>
<th>Study</th>
<th>Was the allocation sequence generated adequately?</th>
<th>Was the allocation of treatment adequately concealed?</th>
<th>Were participants analyzed within the groups they were originally assigned to?</th>
<th>Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches?</th>
<th>Did researchers rule out any impact from a known variable that could cause treatment effect or confound the interpretation of the treatment effect?</th>
<th>Did the study maintain fidelity to the intervention protocol?</th>
<th>If attrition was a concern, were missing data handled appropriately?</th>
<th>Were the outcome assessors blinded to the intervention or exposure status of the participants?</th>
<th>Were the outcomes/exposures assessed using valid and reliable measures, implemented consistently across all study participants?</th>
<th>Were outcomes assessed/exposed using valid and reliable measures, implemented consistently across all study participants?</th>
<th>Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>La Sala, 2015(31)</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td></td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Leandro, 2013(36)</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Legro, 2007(29)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Legro, 2014(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Legro, 2015(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Maher, 2018(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Majumdar, 2013(34)</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Mehrabian, 2012(48)</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Mohammadi Yeganeh, 2018(100)</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Morad, 2012(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Morin-Papunen, 2012(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Mutsaerts, 2016(19)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Nada, 2016(34)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Nahuin, 2011(31)</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Nandi, 2017(30)</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Palomba, 2010(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Palomba, 2011(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Peeraa, 2015(38)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Poruali, 2017(31)</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Qu, 2017(32)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Raghi, 2012(34)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Rahman, 2017(34)</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Rashidi, 2013(31)</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Rashidi, 2015(31)</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Razi, 2013(33)</td>
<td>N</td>
<td>N</td>
<td>U</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Reindollar, 2010(32)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Rubio, 2013(37)</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
</tr>
<tr>
<td>Seckin, 2014(39)</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Selman, 2016(31)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Study</td>
<td>Was the allocation sequence generated adequately? (selection bias)</td>
<td>Was the allocation of treatment adequately concealed? (selection bias)</td>
<td>Were participants analyzed within the groups they were originally assigned to? (selection bias)</td>
<td>Did the design or analysis account for important confounding and modifying variables through matching, stratification, or other approaches? (selection bias)</td>
<td>Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? (performance bias)</td>
<td>If attrition was a concern, were missing data handled appropriately? (attrition bias)</td>
<td>In prospective studies, was the length of follow-up the same between the groups, or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls? (detection bias)</td>
<td>Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias)</td>
<td>Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)</td>
<td>Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)</td>
<td>Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias)</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Shi, 2018[1]*</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Sismanoglu, 2009[2]*</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
</tr>
<tr>
<td>Tehraninejad, 2010[5]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>van Rijswijk, 2017[7]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>van Rumste, 2014[8]</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Weiss, 2018[9]</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Wiser, 2010[10]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Zhu, 2018[12]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Zain, 2009[13]</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Zakherah, 2010[14]</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Zarei, 2016[15]</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>Zheng, 2012[16]</td>
<td>Y</td>
<td>U</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Abbreviations: N=No; U=Unclear; Y=Yes
Figure G-3. Summary of risk of bias assessment for included RCTs

- Was the allocation sequence generated adequately? (selection bias)
- Was the allocation of treatment adequately concealed? (selection bias)
- Were participants analyzed within the groups they were originally assigned to? (selection bias)
- Does the design or analysis control account for important confounding and modifying variables through matching, stratification, multivariable analysis, or other approaches? (selection bias)
- Did researchers rule out any impact from a concurrent intervention or an unintended exposure that might bias results? (performance bias)
- Did the study maintain fidelity to the intervention protocol? (performance bias)
- If attrition was a concern, were missing data handled appropriately? (attrition bias)
- In prospective studies, was the length of follow-up the same between the groups, or in case-control studies, was the time period between the intervention/exposure and outcome the same for cases and controls? (detection bias)
- Were the outcome assessors blinded to the intervention or exposure status of participants? (detection bias)
- Were interventions/exposures assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)
- Were outcomes assessed/defined using valid and reliable measures, implemented consistently across all study participants? (detection bias)
- Were the potential outcomes prespecified by the researchers? Are all prespecified outcomes reported? (reporting bias)
Appendix H. Supplemental Project To Assess the Transparency of Reporting for Trials Evaluating Treatment for Infertility

Authors: Williams JW Jr., Eaton JL, Gierisch JM, Masilamani V, von Isenburg M, Chobot MM

Background and Objectives

Selective reporting can bias estimates of effect, yet methods to detect such biases are limited.\(^1\),\(^2\) Statistical methods for detecting publication bias (e.g., funnel plots, Beggs rank correlation) are underpowered.\(^3\) Comparing outcomes listed under Methods versus those reported under Results in published manuscripts is an expedient but crude method for detecting reporting bias.\(^4\) Another method is to search ClinicalTrials.gov (CT.gov) and (a) compare studies identified there to published studies (to detect publication bias) and (b) compare planned analyses and outcomes reported in CT.gov to those reported in the final publication (to detect reporting bias).\(^4\),\(^5\) The EPC guidance recommends this approach.\(^6\) While conceptually sound, this approach may be labor-intensive, and its utility uncertain.

The overall goal of this project was to evaluate the utility of CT.gov for detecting selective reporting, and to determine the impact of selective reporting on the estimates of treatment effect. A secondary goal was to estimate the person-hours required to complete these analyses.

To accomplish these goals, we used an ongoing review, Management of Infertility, to explore differences between information from published sources included in the review and CT.gov.

Methods

Scope and General Approach

We adopted a pragmatic approach, using methods that could be readily incorporated into future systematic reviews. To maintain feasibility while still applying our methods to a range of interventions, we included KQ 1, KQ 2, and KQ 4 from the Management of Infertility review in this analysis. The KQs are listed below:

KQ 1: What are the comparative safety and effectiveness of available treatment strategies for women with polycystic ovary syndrome (PCOS) who are subfertile/infertile and who wish to become pregnant?

KQ 2: What are the comparative safety and effectiveness of available treatment strategies for women with endometriosis who are subfertile/infertile and who wish to become pregnant?

KQ 4: What are the comparative safety and effectiveness of available treatments for women with tubal or peritoneal factors (e.g., pelvic adhesions) who are subfertile/infertile and who wish to become pregnant?

Searching CT.gov

We searched CT.gov for trials potentially applicable to the KQs with the assistance of our search librarian. Because CT.gov does not use MeSH-based search terms, we adapted the search strategies developed for the *Management of Infertility* review to language appropriate for CT.gov. We conducted two searches, a broad search using the basic interface and a more specific search using the advanced interface in CT.gov. For the broad search, we searched for synonyms for infertility (infertility OR infertile OR subfertility OR subfertile OR sub-fertility OR sub-fertile) in the conditions field and limited our results to interventional studies. For the narrow search, we searched for the same synonyms for infertility in the broader search terms field and combined this with multiple, separate searches for each of the conditions of interest. This narrower search was also limited to interventional studies. Exact search strings used in both searches are given in Appendix A.

Results of the two searches were imported into Excel.

Matching Studies

We matched randomized controlled trials (RCTs) identified in CT.gov with those identified for the *Management of Infertility* review at several levels.

First, we determined whether RCTs reporting a live birth outcome that were included in the *Management of Infertility* review had a matching record in CT.gov. Matching was performed initially using the NCT identifier (NCTID). Our intention was to conduct this matching using a semi-automated process within EndNote. This approach proved infeasible due to inconsistent assignment of NCTIDs to EndNote fields. Thus, all matching was accomplished by manual review. For unmatched studies, we conducted a secondary match using other trial registration numbers and then trial characteristics, including: condition, intervention, sample size, and author/investigator. Matching was performed initially for the broad CT.gov search. We then determined the proportion of matched studies that were not identified by the narrow CT.gov search.

Second, for matched studies (i.e., studies included in the *Management of Infertility* review with a CT.gov record), we abstracted selected variables from the CT.gov record to determine whether key study design variables and reported outcomes matched information in the published manuscript. Variables abstracted were:

- Date of completion
- Number of study arms
- Intervention description
- Study design
- Outcomes measures and results prioritized in the *Management of Infertility* review
- Analysis approach
- Subgroup analyses

Data from CT.gov were compared to published data. For each variable, the result was classified as: matching, discrepant, or possibly discrepant. Discrepant data were defined as cases where information was absent in one source but reported in another, or when the information given in the two sources was contradictory. Discrepancies were summarized narratively.

Third, we screened the unmatched CT.gov citations for potentially eligible completed trials. Eligibility criteria for each KQ are given in Table 1 of the Methods chapter of the main
Management of Infertility review. For potentially eligible studies identified from CT.gov, we used author names and intervention terms to search for a matching publication in PubMed. We classified studies into two groups: (1) potentially eligible completed study without a published manuscript; and (2) potentially eligible completed study with a matching published manuscript that was not identified in the systematic review search.

All matching was limited to studies published since the 2005 International Committee of Medical Journal Editors (ICMJE) policy requiring trial registration. Matching was performed initially by a research assistant, and reviewed by a study investigator. Team members involved in matching piloted the data collection forms and procedures to refine them before full use.

**Estimate of Person-Hours Required To Complete the Project**

EPC staff routinely log the time spent working on projects using project-specific codes. Co-investigators do not log project time routinely. Therefore, our project coordinator sent regular queries to co-investigators asking for estimates of time spent (to nearest 15 minutes) completing project-specific tasks. These estimates were tracked in an Excel spreadsheet. We used the staff logs and co-investigator reports to estimate the total staff time and co-investigator time dedicated to completing project-related activities.

**Impact on Systematic Review Conclusions**

Study conclusions will flow from the strength of evidence (SOE). We used the GRADE framework for evaluating SOE, a framework that includes assessment of risk of bias, consistency, precision, directness, and publication bias. The EPC risk of bias tool explicitly considers reporting bias. Therefore, risk of bias and publication bias are the domains most likely to be affected by supplemental data from CT.gov. In collaboration with authors of the Management of Infertility review, we reviewed the SOE table to determine qualitatively whether study conclusions would change.

**Results**

Results are presented in five sections: (1) concordance between RCTs included in the Management of Infertility review and in CT.gov; (2) studies identified from CT.gov as potentially eligible but not included in the Management of Infertility review; (3) concordance between data from CT.gov and published studies for studies present in both sources; (4) effects of CT.gov results on SOE and review conclusions; and (5) person-hours required to generate these results.

**Concordance Between RCTs Included in the Management of Infertility Review and in CT.gov**

Twenty-four unique RCTs reported live birth as an outcome and were included for KQs 1, 2, and 4 in the Management of Infertility review. The majority of these trials (n=22) were applicable to KQ 1. Of the 24 trials:

- 8 were matched to a CT.gov record by NCTID
- 3 were matched by other trial ID number
- 1 was matched by other criteria (i.e., study characteristics)
- 12 were not matched
All matched studies were confirmed by an investigator. Three preliminary matches based on “other criteria” were not confirmed by study investigators and are included in the 12 unmatched studies above.

Only one-third of the included trials were matched to a CT.gov record using the NCTID, the most reliable and readily applied matching variable. When using all available data, 50% (95% CI, 30 to 50%) of the eligible studies were matched to a CT.gov record.

Studies Identified From CT.gov as Potentially Eligible but Not Included in the Management of Infertility Review

Using broad search criteria, we searched CT.gov for potentially eligible studies. The search yielded 858 registered studies. Of those, 376 were classified as “completed.” The 355 studies published from 2005 forward were reviewed by two study staff, and 94 were flagged as potentially eligible for the Management of Infertility review, with relevance to KQs as follows: KQ1 = 14, KQ 2 = 1, KQ 3 = 69, KQ 4 = 1, KQ 5 = 3, KQ 6 = 1, and multiple KQs = 5.

Of the 16 studies potentially relevant to KQs 1, 2, or 4, 11 had been identified in the Management of Infertility search and included in the review. The other 5 studies were reviewed by an investigator; details are reported in the Table H-1.

Table H-1. Potentially eligible studies not included in the review

<table>
<thead>
<tr>
<th>NCTID</th>
<th>Search Strategy Identifying Trial</th>
<th>CT.gov Completion Date</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCT01675843</td>
<td>Both</td>
<td>March 2012</td>
<td>Potentially eligible; no citation in PubMed</td>
</tr>
<tr>
<td>NCT01679574</td>
<td>Both</td>
<td>January 2012</td>
<td>Potentially eligible; no citation in PubMed</td>
</tr>
<tr>
<td>NCT01894074</td>
<td>Broad</td>
<td>July 2015</td>
<td>Potentially eligible; no citation in PubMed</td>
</tr>
<tr>
<td>NCT00220545</td>
<td>Both</td>
<td>March 2006</td>
<td>Identified in original review search but excluded at title-and-abstract screening stage. Full text reviewed and study included in Management of Infertility review</td>
</tr>
<tr>
<td>NCT01581359</td>
<td>Both</td>
<td>May 2015</td>
<td>Potentially eligible; no citation in PubMed</td>
</tr>
</tbody>
</table>

Only 5 potentially eligible studies were identified across the 3 KQs. Of these, 2 are recently completed trials (2015) and no journal publication was expected. Two trials with a combined sample size of 340 patients were completed more than 3 years ago, indicating potential publication bias. Both of these trials were applicable to KQ 1. One trial was excluded at the title-and-abstract screening phase of the review; upon review of the full text, the study was reclassified as eligible and included in the review.

Concordance Between Data From CT.gov and Published Studies for Studies Present in Both Sources

Study investigators participating in the transparency project abstracted data independently from CT.gov for the 8 studies matched by the NCTID. These data were compared to data abstracted from published data by the Management of Infertility investigators.

Overall, there were no important differences in the study characteristic descriptions between the two sources. Details are described below:
The KQ classification matched for all 8 studies.
The study design and number of study arms matched for all 8 studies.
Of 5 studies reporting the enrolled “n,” 4 were exact matches and 1 had a discrepancy in the estimated enrollment (326) vs. the number enrolled (320). Three studies did not report the sample size in CT.gov and thus were classified as discrepant.
Intervention descriptions were substantially concordant for all 8 studies and thus were classified as matching.
The analytic approach and any plans for subgroup analyses were not addressed in CT.gov for any of the studies. However, subgroup analyses were not reported in the published manuscripts for any of these trials.
The funding sources was classified as matched for 6 studies. Two studies were classified as discrepant: 1 of these was classified as non-government/non-industry from CT.gov and as “not reported” from manuscript, and 1 was classified as non-government/non-industry from CT.gov and as government from the published manuscript.

Outcomes were compared at 2 levels: the outcomes planned from CT.gov to those reported in published manuscripts, and the results reported in CT.gov to those reported in published manuscripts.
- Planned outcomes: 11 outcomes were reported in both sources and classified as matched. Three outcomes reported as planned in CT.gov were not abstracted from manuscripts: quality of life, miscarriage, and live birth. In 4 studies, outcomes reported in published manuscripts were not described in CT.gov: live birth, miscarriage, multiple births, and surgical complications.
- Only 1 of the 8 trials reported results in CT.gov, and these results matched those reported in the manuscript for the single outcome present in both sources.

Effects of CT.gov Results on Strength of Evidence
Overall, data from CT.gov had little impact on the SOE ratings. Using a threshold of 3 years since reported completion, only 2 completed trials were identified from CT.gov that did not have a matching journal publication. Both trials were applicable to KQ 1 and had a combined sample size of 340 patients. Thirty trials (10,718 patients) were included in the SOE rating for KQ 1, and thus these 2 “missing” trials are unlikely to have had a meaningful impact on study results. Similarly, there was little evidence of reporting bias, with only single mismatches for 3 different outcomes between planned outcomes in CT.gov and reported outcomes in published manuscripts.
Person-Hours Required for Data Collection and Analysis

Overall, the project team devoted an estimated 74.5 hours to planning and conducting this study. Data by investigator vs. staff are given in Table H-2.

Table H-2. Person-hours required, investigators vs. staff

<table>
<thead>
<tr>
<th>Name</th>
<th>Administrative (meetings, etc)</th>
<th>Planning/designing</th>
<th>Running searches/abstracting data</th>
<th>Synthesizing data/writing</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Investigator</td>
<td>7</td>
<td>9</td>
<td>10.75</td>
<td>7.75</td>
<td>34.5</td>
</tr>
<tr>
<td>EPC Staff</td>
<td>23</td>
<td>0</td>
<td>14</td>
<td>3</td>
<td>40</td>
</tr>
<tr>
<td>Totals</td>
<td>30</td>
<td>9</td>
<td>24.75</td>
<td>10.75</td>
<td>74.5</td>
</tr>
</tbody>
</table>

Discussion

This substudy found that CT.gov has important limitations for identifying selective reporting. Only one-third of the studies included in the Management of Infertility review were matched to a CT.gov record based on NCTID, and only 1 of those studies reported results in CT.gov. In addition, there were few discrepancies between planned outcomes reported in CT.gov and those reported in published manuscripts. A careful search and inspection of CT.gov for potentially eligible studies not identified by the review team yielded only 2 studies without a publication and 1 study incorrectly excluded at the title-and-abstract screening stage. These data had no impact on the SOE ratings or study conclusions, but required substantial person-hours to generate.

It is possible that CT.gov will mature into a more useful resource for the purpose of identifying selective reporting. Using data from CT.gov for the dates of trial registration compared to conduct of the study, it is clear that some studies were registered retrospectively. Prospective registration may yield more complete records and more informative data. However, it is likely that changes to CT.gov will be required for this database to serve as a useful source for identifying selective reporting.

At present, these results do not support the routine use of CT.gov to evaluate selective reporting. However, our study examined a small set of interventions for a single condition (infertility) and included a relatively small set of trials. Additional studies are needed before definitive conclusions can be drawn about the utility of CT.gov for detecting selective reporting. If changes to CT.gov were made to facilitate its use for this purpose, other resources could improve efficiency, including: a customized EndNote filter for importing CT.gov results, a standard methodology to guide investigators, and additional data on the activities that can be reliably completed by study staff versus investigators.

References to Appendix H


