Evidence-based Practice Center Systematic Review Protocol

Project Title: Management of Infertility

I. Background and Objectives for the Systematic Review

"Abnormal" fertility is typically defined based on the duration of time a couple has been attempting to achieve pregnancy. "Infertility" has traditionally been defined as failure to achieve pregnancy after 12 months of regular unprotected intercourse with the same partner; because as many as half of these couples will conceive without intervention over the next 12-24 months, the term "subfertility" is also sometimes used.[1] Self-reported infertility, 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). An additional 5 percent report difficulty in carrying a pregnancy to term.[2] In one population-based study, approximately 10 percent of pregnant women reported receiving infertility treatment, with 29 percent reporting use of fertility-enhancing medications; 21 percent use of assisted reproductive technology (ART), including *in vitro* fertilization (IVF); 15 percent artificial insemination with fertility-enhancing drugs; and 23 percent other treatments, including surgery.[3] Other estimates of the prevalence of infertility treatment are similar.[4-8]

The most common demographic factor associated with female infertility/impaired fecundity is "advanced" 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% of cycles.[11] Other common causes of female infertility include 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]

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]

Available Treatments: 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 IVF and intra-cytoplasmic sperm injection (ICSI).[16,17] Some of these treatments (surgical management of endometriosis, or ovarian "drilling" for PCOS) are less commonly used in current U.S. practice. Treatment options for male factor infertility include medical treatment of a diagnosed endocrinopathy or other condition affecting sperm production, empiric treatments with hormonal or other agents, surgical management of varicocele, intrauterine insemination, IVF, and ICSI.[18] Options appropriate for some diagnoses (e.g.,

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

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. Trade-offs between outcomes (particularly multiple gestations), time to pregnancy, and out-of-pocket costs might be different between the two strategies even if cumulative live birth rates were identical.

Outcomes of Interest: 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[19-23] However, there is a growing consensus that the most important patient centered outcome should be live birth.[22,23]

Adverse outcomes of treatment: Different treatments also carry different safety risks. There are known short-term risks such as ovarian hyperstimulation syndrome (OHSS). Although the majority of 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 still some uncertainty surrounding some outcomes in subgroups of patients.[24-26]

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.[19,21,27,28] There may also be direct effects of some treatments that have unclear implications for long-term health in children born after these treatments—for example, the possibility that epigenetic changes may lead to increased risk of some disorders later in life.[29,30]

Infertility clearly has an emotional impact,[14,31,32] and the comparative effects of infertility treatments on quality of life are an important consideration.

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. There is evidence that 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]

Infertility treatment is a topic where there 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

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

adverse outcomes, and in some cases preferences for process as well as outcome. In addition, time is an important consideration, particularly for women 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 other 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 transfer[48] and use of fresh vs. 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]

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, it is limited by (a) use of the ART cycle (rather than the individual patient) as the unit of analysis, and (b) lack of long-term follow-up data for individual patients;[38] there is also some concern about underreporting of some adverse outcomes.[54] 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. The 2008 AHRQ Evidence Report on "Effectiveness of ART"[55,56] found that approximately 80 percent of the 478 included studies were performed outside the United States. The majority of RCTs were not designed to detect differences in pregnancy and live birth rates; reporting of delivery rates and obstetric outcomes was unusual. 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. With advances in research design and analysis, and establishment of an active NIH-funded clinical network, some of these gaps may have been filled. In addition, the 2008 report was focused on outcomes of specific treatments, rather than organized around the underlying diagnosed cause of infertility, so some relevant questions of interest to patients and other stakeholders (such as the comparative effectiveness of ART vs. tubal re-anastomosis for patients seeking reversal of sterilization) were not addressed.

II. The Key Questions

The draft key questions (KQs) developed during Topic Refinement were available for public comment from June 9, 2015 to June 29, 2015. Overall, the comments affirmed our planned

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

approach. Specific suggestions fell into three broad categories: (a) suggestions for additional specific subgroups, interventions, and outcomes to be included, (b) suggestions which affected the scope of the review (for example, limiting the review to U.S.-based studies only because of differences in practice and health systems, conducting primary data analyses, or expanding the range of costs to be considered), and (c) statements/recommendations for care which assumed that the answers to the KQs were already known. Specific recommendations concerning PICOTS were added to the existing list (*indicated in italics*). The overall scope of the review was not changed in response to public suggestions, but the very relevant issues raised by these comments/suggestions will be addressed in the discussion of the review's findings. There were no other significant changes to the KQs or proposed methods.

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?

a. 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 subfertile/infertile and who wish to become pregnant?

a. 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 subfertile/infertile for unknown reasons and who wish to become pregnant?

a. 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 treatments for women with tubal factor or *peritoneal factor (e.g., pelvic adhesions)* infertility who are subfertile/infertile and who wish to become pregnant?

a. 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 treatments for couples with male factor infertility and no evidence of an underlying diagnosis associated with infertility in the female partner?

a. 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?

a. For female oocyte donors:

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

i. 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)?

ii. Do short- and long-term outcomes differ among different induction/retrieval protocols?

b. For male semen donors:

i. Are there long-term health, quality-of-life, or other adverse outcomes associated with donation?

KQ 1:

- **Population:**
  - 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 (alternate definitions may be appropriate), and diagnosed PCOS. Subpopulations of interest include groups differing in age; race/ethnicity; obesity/BMI; ovarian reserve; history of prior treatments; primary vs. secondary infertility; *maternal parity*; diagnostic criteria/evaluation (e.*g., WHO categories*); *insurance status (particularly coverage of infertility diagnosis and treatment);* and presence or absence of male factor infertility, other female causes of infertility, or *common comorbidities such as hypertension and diabetes.*
- **Interventions:**
  - 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
- **Comparators:**
  - 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).
- **Outcomes:** (Note that these outcomes 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 (both cumulative and per cycle)
    - Live singleton birth
      - Gestational age at birth/preterm birth (subcategorized as preterm/very preterm)
    - Live multiple birth

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

        - Gestational age at birth/preterm birth (subcategorized as preterm/very preterm)
    - Pregnancies (both cumulative and per cycle)
        - "Biochemical" pregnancy (positive human chorionic gonadotropin [hCG])
        - Presence of gestational sac
        - Positive fetal heart rate
    - Pregnancy complications
        - Multiple births (and associated complications)
        - Preeclampsia
        - Ectopic pregnancies
        - Miscarriage
        - *Premature rupture of membranes*
        - *Gestational diabetes*
        - *Placental abnormalities (placental abruption, placenta previa)*
        - *Mode of delivery (spontaneous vaginal delivery, instrumental vaginal delivery, cesarean section)*
    - Neonatal outcomes
        - Death
        - Birthweight (categorized as low birthweight/normal birthweight)
        - Congenital anomalies
    - Time to pregnancy
        - Calendar time (months)
        - Number of cycles
    - Quality of life/psychological impact (short- and long-term)
        - Maternal
        - Paternal
        - Couple
    - Costs
        - Patient
        - Health system
        - Societal
    - Short-term adverse effects of treatments
        - OHSS

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

        - Surgical complications
        - Adverse effects of treatments (e.g., for PCOS patients, gastrointestinal [GI] symptoms for metformin, hot flashes for clomiphene)
    - Long-term outcomes (child)
        - Neurodevelopmental/other issues related to prematurity
        - Specific issues related to infertility treatment (epigenetic changes*, sex chromosomal abnormalities*, etc.)
        - Cancer (all types)
        - Others
    - Long-term outcomes (maternal)
        - Cancer
        - Subsequent fertility
        - Age at menopause
        - Others

- **Timing:**
    - 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

- **Settings:**
    - Subspecialty practice (infertility specialist)
    - General gynecology practice
    - Family practice/general internist/nurse practitioner/other non-gynecologist primary care provider
    - *United States vs. non-U.S.*

KQ 2:

- **Population:**
    - 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 (alternate definitions may be appropriate), and diagnosed endometriosis. 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;* diagnostic criteria/evaluation; stage of endometriosis; presence of endometriomia (ovarian cyst with endometrial tissue); and presence

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

or absence of male factor infertility, other female causes of infertility, or common comorbidities such as hypertension and diabetes.

- **Interventions:**
    - 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, superovulation with gonadotropins with or without IUI, ART (IVF and ICSI) with patient and donor oocytes
- **Comparators:**
    - Either be direct between two alternatives (e.g., surgery vs. gonadotropin-releasing hormone [GnRH] agonists/antagonists), or timing/sequence of interventions (e.g., ovulation induction/IUI followed by ART if unsuccessful vs. proceeding directly to ART).
- **Outcomes:** Same as for KQ 1
- **Timing:** Same as for KQ 1
- **Settings:** Same as for KQ 1

KQ 3:

- **Population:**
    - 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 (alternate definitions may be appropriate), and no other diagnosed cause of subfertility/infertility. 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;* diagnostic criteria/evaluation; and presence or absence of common comorbidities such as hypertension and diabetes. Women without male partners (single women or lesbian couples) are also a subgroup of interest, particularly for long-term outcomes.
- **Interventions:**
    - Timed intercourse with various technologies, oral ovulation induction agents (e.g., clomiphene citrate), ovulation induction with gonadotropins with and without IUI, ART (IVF and ICSI) with patient and donor oocytes, watchful waiting
- **Comparators:**
    - 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).
- **Outcomes:** Same as for KQ 1
- **Timing:** Same as for KQ 1

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

- **Settings:** Same as for KQ 1

KQ 4:

- **Population:**
    - 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 (alternate definitions may be appropriate), and identified tubal *or peritoneal* disease potentially amenable to surgical interventions (hydrosalpinx, unilateral occlusion, prior tubal sterilization). Subpopulations of interest include groups differing in age, race/ethnicity, obesity/BMI, history of prior treatments, *anatomic* cause of tubal occlusion (e.g., prior sterilization vs. adhesions), *maternal parity, insurance status,* and primary vs. secondary infertility.
- **Interventions:**
    - Surgical repair, ART (IVF and ICSI) with patient and donor oocytes
- **Comparators:**
    - Other active interventions (including combinations of therapy such as surgical removal of hydrosalpinx followed by ART)
- **Outcomes:** Same as for KQ 1
- **Timing:** Same as for KQ 1
- **Settings:** Same as for KQ 1

KQ 5:

- **Population:**
    - 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 (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.
- **Interventions:**
    - 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
- **Comparators:**
    - Any other active intervention
- **Outcomes:** Same as for KQ 1
- **Timing:** Same as for KQ 1

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

- **Settings:**
    - Same as for KQ 1, with the addition of male reproductive medicine specialist/urologist.

KQ 6:

- **Population:**
    - Women of reproductive age (18-44) who are potential donors of oocytes for ART, and males donating semen for intrauterine insemination or ART
- **Interventions (women):**
    - Pre-donation testing strategies; controlled ovarian hyperstimulation with gonadotropins using different induction/retrieval protocols
- **Interventions (men)**
    - Semen donation
- **Comparators (women):**
    - Any other active intervention, or women who are NOT undergoing ovulation induction for oocyte donation
- **Comparators (men)**
    - Men who do not donate semen
- **Outcomes (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
- **Outcomes (men):**
    - Quality-of-life outcomes
    - Short- or long-term health outcomes
- **Timing :**
    - Short term:
        - From time of beginning donation process to 12 months after donation
    - Long-term:

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

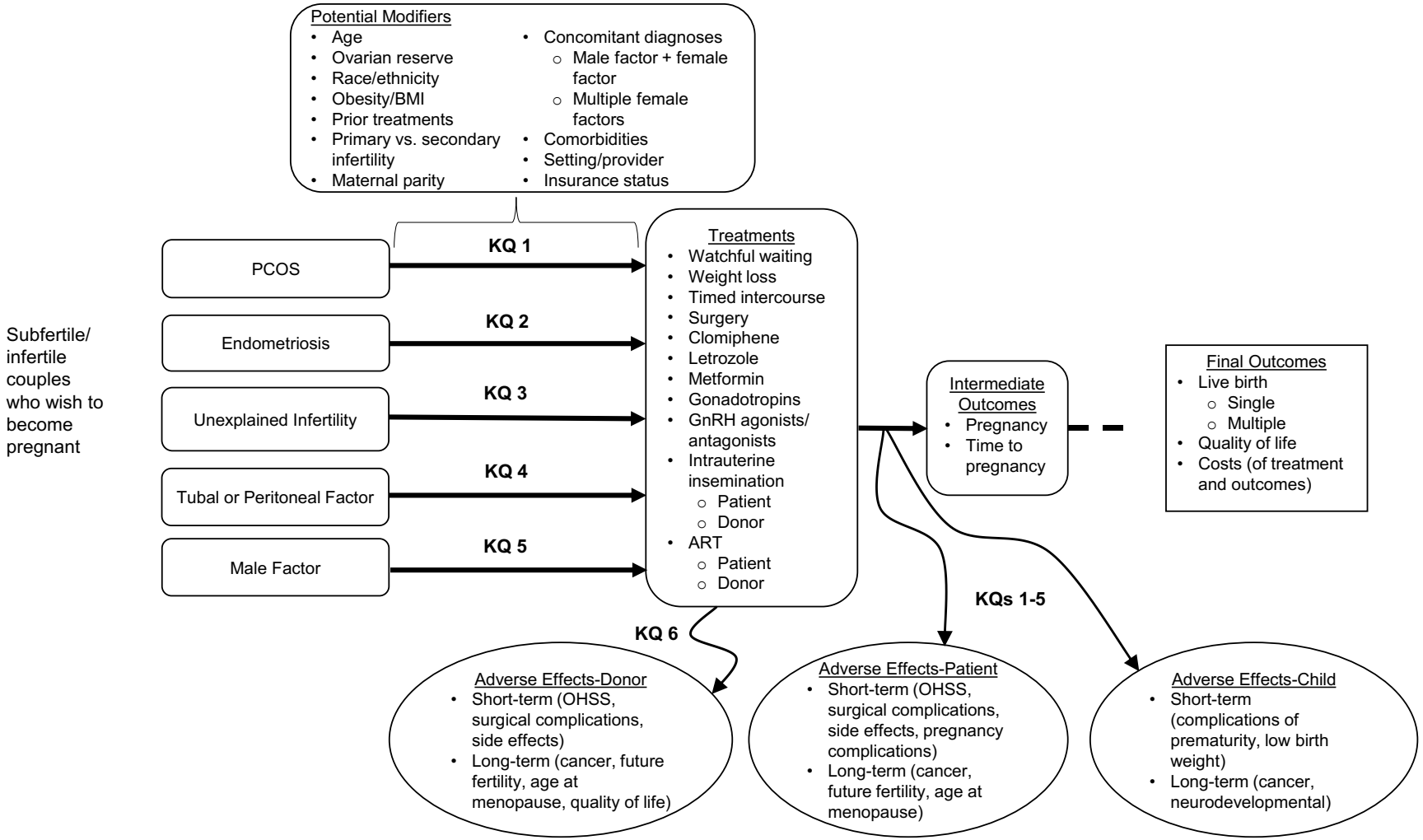
        - 12 months or more from time of first donation
- **Setting:** Subspecialty practice (infertility specialist)

III. Analytic Framework

The analytic framework depicts the key questions within the context of the population, interventions, comparators, outcomes, timings, and settings (PICOTS) described in the previous section. The figure illustrates how a wide range of treatments for infertility may result in intermediate outcomes such as pregnancy or time to pregnancy and/or final outcomes such as live birth (single or multiple), quality of life, or costs in couples with different underlying causes of infertility. A separate key question 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.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

Figure 1. Analytic Framework



Abbreviations: ART=assisted reproductive technology; BMI=body mass index; GnRH=gonadotropin-releasing hormone; KQ(s)=key question(s); OHSS=ovarian hyperstimulation syndrome; PCOS=polycystic ovary syndrome

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

IV. Methods

In developing this comprehensive review, we will apply the rules of evidence and evaluation of strength of evidence recommended by the Agency for Healthcare Research and Quality (AHRQ) in its *Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (hereafter referred to as the *Methods Guide*).[57] We will solicit feedback regarding conduct of the work (such as development of search strategies and identifying outcomes of key importance) from the Task Order Officer and the Technical Expert Panel. We will follow the methodology recommended to the Evidence-based Practice Centers for literature search strategies, inclusion/exclusion of studies in our review, abstract screening, data abstraction and management, assessment of methodological quality of individual studies, data synthesis, and grading of evidence for each KQ.

Table 1. Inclusion and exclusion criteria

PICOTS Element	Inclusion Criteria	Exclusion Criteria

Populations	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 (alternate definitions may be appropriate), and
	  ○ diagnosed PCOS [KQ 1]
	  ○ diagnosed endometriosis [KQ 2]
	  ○ no other diagnosed cause of subfertility/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:
	• 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 (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
		Individuals younger than 18 or 45 and older

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

PICOTS Element	Inclusion Criteria	Exclusion Criteria
	common comorbidities such as hypertension and diabetes. KQ 6: • Women of reproductive age (18-44) who are potential donors of oocytes for ART, and males donating semen for intrauterine insemination or ART	
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 KQ 2: 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 KQ 3: 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 KQ 4: Surgical repair, ART (IVF and ICSI) with patient and donor oocytes KQ 5: 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 KQ 6: 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) KQ 2: Either direct between two alternatives (e.g., surgery vs. gonadotropin-releasing hormone [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	KQ 1-5: • Live birth (both cumulative and per cycle)	KQ1-5: We will exclude studies which only report post-intervention

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

PICOTS Element	Inclusion Criteria	Exclusion Criteria
	○ Live singleton birth ▪ Gestational age at birth/preterm birth (subcategorized as preterm/very preterm) ○ Live multiple birth ▪ Gestational age at birth/preterm birth (subcategorized as preterm/very preterm) • Pregnancies (both cumulative and per cycle) ○ "Biochemical" pregnancy (positive human chorionic gonadotropin [hCG]) ○ Gestational sac ○ Positive fetal heart rate • Pregnancy complications ○ Multiple births (and associated complications) ○ Preeclampsia ○ Ectopic pregnancies ○ Miscarriage ○ Premature rupture of membranes ○ Gestational diabetes ○ Placental abnormalities (placental abruption, placenta previa) ○ Mode of delivery (spontaneous vaginal delivery, instrumental vaginal delivery, cesarean section) • Neonatal outcomes ○ Death ○ Birthweight (categorized as low birthweight/normal birthweight) ○ Congenital anomalies • Time to pregnancy ○ Calendar time (months) ○ Number of cycles • Quality of life/psychological impact (short- and long-term) ○ Maternal ○ Paternal ○ Couple • Costs ○ Patient ○ Health system ○ Societal • Short-term adverse effects of treatments ○ OHSS	pregnancy rates which do not report another outcome of interest (e.g., live birth or complications of treatment). Pregnancy rates will only be abstracted in studies which live births were also reported. We will record papers excluded on this basis.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

PICOTS Element	Inclusion Criteria	Exclusion Criteria
	○ Surgical complications ○ Adverse effects of treatments (e.g., for PCOS patients, gastrointestinal [GI] symptoms for metformin, hot flashes for clomiphene) • Long-term outcomes (child) ○ Neurodevelopmental/other issues related to prematurity ○ Specific issues related to infertility treatment (epigenetic changes, sex chromosomal abnormalities, etc.) ○ Cancer (all types) ○ Others • Long-term outcomes (maternal) ○ Cancer ○ Subsequent fertility ○ Age at menopause ○ Others 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]	

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

PICOTS Element	Inclusion Criteria	Exclusion Criteria
Study design	• Original data, including systematic reviews and meta-analyses • RCTs, prospective and retrospective observational studies with comparator; for test characteristics, cross-sectional studies are acceptable if they include patients with diagnostic uncertainty and direct comparison of test results with an appropriate reference standard • RCTs: All sample sizes • Observational studies: sample size ≥100 subjects	Editorials, nonsystematic reviews, abstracts only, letters, case series, case reports
Publications	• English-language only • Published January 1, 2007 to present	Given the high volume of literature available in English-language publications, non-English articles will be excluded[a]

[a]It is the opinion of the investigators that the resources required to translate non-English articles would not be justified by the low potential likelihood of identifying relevant data unavailable from English-language sources.
Abbreviations: ART = assisted reproductive technology; BMI = body mass index; GI = gastrointestinal; GnRH = gonadotropin-releasing hormone; hCG = human chorionic gonadotropin; ICSI = intra-cytoplasmic sperm injection; IUI = intrauterine insemination; IVF = *in vitro* fertilization; KQ(s) = key question(s); OHSS = ovarian hyperstimulation syndrome; PCOS = polycystic ovary syndrome; PICOTS = Populations, Interventions, Comparators, Outcomes, Timing, Settings; RCTs = randomized controlled trials

Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

To identify relevant published literature, we will search PubMed®, Embase®, and the Cochrane Database of Systematic Reviews, limiting the search to studies conducted in adults and published from January 1, 2007, to the present, based on the cut-off date from the previous evidence report on ART[58] and input from Key Informants, who felt that the Cochrane reviews would identify older relevant high-quality studies, particularly evidence from RCTs, while primary studies and other systematic reviews published subsequent to the ART report would identify studies most relevant to current practice in infertility. Our proposed search strategy for PubMed® is provided in Table 2; this strategy will be adapted as appropriate for searching the other databases. Where possible, we will use existing validated search filters (such as the Clinical Queries Filters in PubMed®). An experienced search librarian will guide all searches. We will supplement the electronic searches with a manual search of citations from a set of key primary and review articles. The reference list for identified pivotal articles will be manually hand-searched and cross-referenced against our database, and additional relevant manuscripts will be retrieved. All citations will be imported into an electronic bibliographical database (EndNote® Version X4; Thomson Reuters, Philadelphia, PA).

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

Table 2: PubMed Search Strategy

Set #	Terms
#1	"Infertility"[Mesh] OR "Anovulation"[Mesh] OR “infertility”[tiab] OR “infertile”[tiab] OR “subfertility”[tiab] OR “subfertile”[tiab] OR “sub-fertility”[tiab] OR “sub-fertile”[tiab] OR “anovulation”[tiab] OR “aspermia”[tiab] OR “asthenozoospermia”[tiab] OR “azoospermia”[tiab] OR “oligospermia”[tiab] OR “sertoli cell-only syndrome”[tiab]
#2	"Reproductive Techniques, Assisted"[Mesh] OR "Polycystic Ovary Syndrome/therapy"[Mesh] OR "Endometriosis/therapy"[Mesh] OR "Nutrition Therapy"[Mesh] OR "Weight Loss"[Mesh] OR "Exercise"[Mesh] OR "Exercise Therapy"[Mesh] OR "Fertility Agents"[Mesh] OR "Clomiphene"[Mesh] OR "Gonadotropin-Releasing Hormone"[Mesh] OR "Metformin"[Mesh] OR "Hormone Antagonists"[Mesh] OR "Gonadotropins"[Mesh] OR “Watchful Waiting"[Mesh] OR “Natural Family Planning Methods “[MeSH] OR “Ovulation Detection”[MeSH] OR "Fallopian Tubes/surgery"[Mesh] OR "Fallopian Tube Diseases/surgery"[Mesh] OR "Gynecologic Surgical Procedures"[Mesh] OR “Arginine/therapeutic use”[MeSH] OR “Aspartic Acid/therapeutic use”[MeSH] OR “Citrulline/therapeutic use”[MeSH] OR “Flavonoids/therapeutic use”[MeSH] OR “Adrenal Cortex Hormones”[MeSH] OR “Ejaculatory Ducts/therapy”[MeSH] OR "Varicocele/surgery"[Mesh] OR “Laser Therapy”[MeSH] OR “Dexamethasone”[MeSH] OR “Vasovasostomy”[MeSH] OR “Urofollitropin”[MeSH] OR “electrocoagulation”[MeSH] OR “Preimplantation Diagnosis”[MeSH] OR "Insemination"[Mesh] OR "Uterine Diseases/surgery"[Mesh] OR "Reproductive Techniques"[Mesh:NoExp] OR "Ovulation Prediction"[Mesh] OR "Genetic Testing"[Mesh] OR "letrozole"[Supplementary Concept] OR “cetrorelix”[Supplementary Concept] OR “ganirelix”[Supplementary Concept] OR “follitropin beta”[Supplementary Concept] OR “follitropin alfa”[Supplementary Concept] OR “Crinone”[Supplementary Concept] OR “Ovidrel”[Supplementary Concept] OR “reproductive techniques”[tiab] OR “reproductive technology”[tiab] OR “reproductive technique”[tiab] OR “reproductive technologies”[tiab] OR “assisted reproductive”[tiab] OR “ivf”[tiab] OR “in vitro”[tiab] OR invitro[tiab] OR “sperm injection”[tiab] OR “ICSI”[tiab] OR “IUI”[tiab] OR “intrauterine insemination”[tiab] OR “intrauterine implantation”[tiab] OR “embryo transfer”[tiab] OR “artificial insemination”[tiab] OR “assisted pregnancy”[tiab] OR “assisted reproduction”[tiab] OR “ovulation induction”[tiab] OR “ovarian stimulation”[tiab] OR “ovarian hyperstimulation”[tiab] OR “clomiphene”[tiab] OR “serophene”[tiab] OR “clomiphene citrate”[tiab] OR “letrozole”[tiab] OR “metformin”[tiab] OR “gonadotropins”[tiab] OR “gonadotropin releasing hormone”[tiab] OR “hormone antagonists”[tiab] OR “menotropins”[tiab] OR “menopur”[tiab] OR “repronex”[tiab] OR “goserelin”[tiab] OR “Zoladex”[tiab] OR “leuprolide”[tiab] OR “Lupron”[tiab] OR “nafarelin”[tiab] OR “Synarel”[tiab] OR “cetrorelix”[tiab] OR “Cetrotide”[tiab] OR “degarelix”[tiab] OR “Firmagon”[tiab] OR “ganirelix”[tiab] OR “antagon”[tiab] OR exercise[tiab] OR diet[tiab] OR “weight loss”[tiab] OR “natural family planning”[tiab] OR “timed intercourse”[tiab] OR “Billings”[tiab] OR “Creighton”[tiab] OR “rhythm method”[tiab] OR “standard days method”[tiab] OR “calendar method”[tiab] OR “basal body temperature method”[tiab] OR “hysteroscopy”[tiab] OR “hysteroscopic”[tiab] OR “microhysteroscopy”[tiab] OR “microhysteroscopic”[tiab] OR “ovarian drilling”[tiab] OR “donor oocytes”[tiab] OR “oocyte retrieval”[tiab] OR “sperm donation”[tiab] OR “sperm donor”[tiab] OR “semen donation”[tiab] OR “semen donor”[tiab] OR “sperm extraction”[tiab] OR “sperm retrieval”[tiab] OR “sperm aspiration”[tiab] OR “tesa”[tiab] OR “micro tese”[tiab] OR “mesa”[tiab] OR “pesa”[tiab] OR “ejaculatory duct resection”[tiab] OR “recombinant human follicle stimulating hormone”[tiab] OR “rhFSH”[tiab] OR “rFSH”[tiab] OR “hormone therapy”[tiab] OR “laser vaporization”[tiab] OR “laser vaporisation”[tiab] OR “dexamethasone”[tiab] OR “vasectomy reversal”[tiab] OR “sterilization reversal”[tiab] OR “superovulation”[tiab] OR “follistim”[tiab] OR “Gonal F”[tiab] OR “Gonal-F”[tiab] OR “Bravelle”[tiab] OR “crinone”[tiab] OR “endometrim”[tiab] OR “prometrium”[tiab] OR “fulguration”[tiab] OR “endometriosis excision”[tiab] OR “endometrioma excision”[tiab] OR “ovarian cystectomy”[tiab] OR “tubal ligation reversal”[tiab] OR “tubal cannulation”[tiab] OR

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

Set #	Terms
	“therapeutic donor insemination”[tiab] OR “ovulation prediction”[tiab] OR “ovidrel”[tiab] OR “assisted hatching”[tiab] OR “preimplantation diagnosis”[tiab] OR “preimplantation genetic diagnosis”[tiab] OR “preimplantation screening”[tiab] OR “preimplantation genetic screening”[tiab] OR “preimplantation testing”[tiab] OR “preimplantation genetic testing”[tiab]
#3	(randomized controlled trial[pt] OR controlled clinical trial[pt] OR randomized[tiab] OR randomised[tiab] OR randomization[tiab] OR randomisation[tiab] OR randomly[tiab] OR Clinical trial[pt] OR “clinical trial”[tiab] OR “clinical trials”[tiab] OR "evaluation studies"[Publication Type] OR "evaluation studies as topic"[MeSH Terms] OR "evaluation study"[tiab] OR evaluation studies[tiab] OR "intervention studies"[MeSH Terms] OR "intervention study"[tiab] OR "intervention studies"[tiab] OR "case-control studies"[MeSH Terms] OR "case-control"[tiab] OR "cohort studies"[MeSH Terms] OR cohort[tiab] OR "longitudinal studies"[MeSH Terms] OR "longitudinal”[tiab] OR longitudinally[tiab] OR "prospective"[tiab] OR prospectively[tiab] OR "retrospective studies"[MeSH Terms] OR "retrospective"[tiab] OR "follow up"[tiab] OR "comparative study"[Publication Type] OR "comparative study"[tiab] OR systematic[subset] OR "meta-analysis"[Publication Type] OR "meta-analysis as topic"[MeSH Terms] OR "meta-analysis"[tiab] OR "meta-analyses"[tiab]) NOT (Editorial[ptyp] OR Letter[ptyp] OR Case Reports[ptyp] OR Comment[ptyp]) NOT (animals[mh] NOT humans[mh])
#4	#1 AND #2 AND #3
	Dates: 2008/01/01 - present
	Limit: English

As a mechanism to ascertain publication bias in recent studies, we will search ClinicalTrials.gov to identify completed but unpublished studies (we will also explore the possibility of publication bias specifically in our quantitative synthesis of the included literature through meta-analysis techniques). While the draft report is under peer review, we will update the search and include any eligible studies identified either during that search or through peer or public reviews in the final report.

We will use several approaches to identifying relevant gray literature, including notification to stakeholders, including drug and device manufacturers, of requests to submit scientific information packets and a search of U.S. Food and Drug Administration (FDA) device registration studies and new drug applications. We will also search study registries for relevant articles from completed studies. Gray literature databases will include ClinicalTrials.gov and the World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP) search portal.

For citations retrieved from MEDLINE, Embase, and the Cochrane Database of Systematic Reviews, two reviewers using prespecified inclusion/exclusion criteria will review titles and abstracts for potential relevance to the research questions. Articles included by either reviewer will undergo full-text screening. At the full-text screening stage, two independent reviewers must agree on a final inclusion/exclusion decision. Articles meeting eligibility criteria (see Table, above) will be included for data abstraction. All results will be tracked using the DistillerSR data synthesis software program (Evidence Partners Inc., Manotick, ON, Canada).

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

Data Abstraction and Data Management

The research team will create data abstraction forms for the KQs that will be programmed in the DistillerSR software. Based on their clinical and methodological expertise, a pair of researchers will be assigned to abstract data from each of the eligible articles. One researcher will abstract the data, and the second will over-read the article and the accompanying abstraction to check for accuracy and completeness. Disagreements will be resolved by consensus or by obtaining a third reviewer's opinion if consensus cannot be reached. Guidance documents will be drafted and provided to the researchers to aid both reproducibility and standardization of data collection.

We will design the data abstraction forms for this project 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 will pay particular attention to describing the details of the treatment (e.g., for comparisons of 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 versus observational) that may be related to outcomes. In addition, we will describe comparators carefully, as treatment standards may have changed during the period covered by the review. The safety outcomes will be 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 the Methods Guide,[57] will also be abstracted. Before they are used, abstraction-form templates will be pilot-tested with a sample of included articles to ensure that all relevant data elements are captured and that there is consistency/reproducibility between abstractors. Forms will be revised as necessary before full abstraction of all included articles. Final abstracted data will be uploaded to SRDR per EPC requirements.

Assessment of Methodological Risk of Bias of Individual Studies

We will assess 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.[57] Briefly, we will rate each study as being of good, fair, or poor quality based on its adherence to well-accepted standard methodologies. For all studies, the overall study quality will be assessed as follows:

- Good (low risk of bias). 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.
- Fair. 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

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

may have been missing information, making it difficult to assess limitations and potential problems.

- Poor (high risk of bias). 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.

The grading will be outcome-specific such that a given study that analyzes its primary outcome well but did an incomplete analysis of a secondary outcome would be assigned a different quality grade for each of the two outcomes. Studies of different designs will be graded within the context of their respective designs. Thus, RCTs will be graded good, fair, or poor, and observational studies will separately be graded good, fair, or poor.

Data Synthesis

We will begin by summarizing key features of the included studies for each KQ. To the degree that data are available, we will abstract information on study design; patient characteristics; clinical settings; interventions; and intermediate, final, and adverse event outcomes.

We will then determine 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, conceptual homogeneity of the studies, completeness of the reporting of results, and the adequacy and completeness of any existing meta-analyses. Because there are a large number of existing systematic reviews for this topic, particularly from the Cochrane group, we will consider these results using suggested guidance from the Methods Guide chapter on integrated bodies of evidence,[59] as outlined in more detail below. As recommended by the Guide, judgments about the benefit of performing a new quantitative estimate will be based on an assessment of the existing strength of evidence (using the domains of study limitations, consistency, precision, directness and reporting bias) and a judgment about the degree to which a new quantitative synthesis would change conclusions about benefit harm/trade-offs, assessment of strength of evidence, substantially improve the precision of the estimate, or provide a more up-to-date estimate reflecting current practice.

When a meta-analysis is appropriate, we will use random-effects models to synthesize the available evidence quantitatively. We will test for heterogeneity using graphical displays and test statistics (Q and I^2 statistics), while recognizing that the ability of statistical methods to detect heterogeneity may be limited. For comparison, we will also perform fixed-effect meta-analyses. We will present summary estimates, standard errors, and confidence intervals. We anticipate that intervention effects may be heterogeneous. We hypothesize that the methodological quality of individual studies, study type, the characteristics of the comparator, and patients' underlying clinical presentation will be associated with the intervention effects. If there are sufficient studies, we will perform subgroup analyses and/or meta-regression analyses to examine these hypotheses. We will perform quantitative and qualitative syntheses separately by study type and discuss their consistency qualitatively.

For a decision analysis or simulation model, feasibility will be based on a judgment about the degree to which such an analysis will provide additional insight into the key questions based

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

on the available evidence—for example, a stochastic simulation of the likelihoods of live birth, multiple gestation, preterm delivery, and complications of pre-term birth over several cycles of two different ART protocols based on results of a meta-analysis of relevant trials would give insight into the existing degree of certainty about the benefit-harm trade-off associated with each protocol, which would inform future research prioritization.[60]

When a decision analysis is appropriate, we will adapt existing models of conception and pregnancy developed by Dr. Myers to accommodate specific questions.[61] We will use suggested guidance on the use of simulation models in EPC reports as developed by the Brown EPC.[62]

Grading the Strength of Evidence (SOE) for Major Comparisons and Outcomes

We will grade the strength of evidence for each outcome assessed; thus, the strength of evidence for two separate outcomes in a given study may be graded differently. The strength of evidence will be assessed using the approach described in the AHRQ's Methods Guide.[57,63] In brief, the approach requires assessment of five domains: study limitations (previously named risk of bias), consistency, directness, precision, and reporting bias, which includes publication bias, outcome reporting, and analysis reporting bias, as described in detail above. Additional domains to be used when appropriate (most relevant to observational studies) are dose-response association, impact of plausible residual confounders, and strength of association (magnitude of effect). When the body of evidence for a particular outcome includes both RCTs and observational studies, we will grade each study type separately using design-specific criteria. In considering the overall strength of the entire body of evidence, we will consider the extent to which the observational evidence is consistent with RCT data, particularly with regards to direction and magnitude of effect. Because of the risk of unmeasured confounding, observational studies would generally not contribute to estimates of the magnitude of effect, and judgment about the precision of the effect, when RCT data is available. If there are 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 lead to increased uncertainty about the magnitude and precision of any treatment effect.[64] These domains will be considered qualitatively, and a summary rating of high, moderate, or low strength of evidence will be assigned for each outcome after discussion by two reviewers. In some cases, high, moderate, or low ratings will be impossible or imprudent to make, for example, when no evidence is available or when evidence on the outcome is too weak, sparse, or inconsistent to permit any conclusion to be drawn. In these situations, a grade of "insufficient" will be assigned. This four-level rating scale consists of the following definitions:

- High—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—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—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

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

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—We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

As noted above, there has already been a large body of systematic reviews, some with meta-analyses, in this area. We will use the recommendations outlined in the Methods Guide chapter on integrating existing systematic reviews in incorporating this body of evidence into our review. Briefly, we will confirm that a given paper is a systematic review by requiring that the review include an explicit and adequate search, application of predefined eligibility criteria to select studies, risk of bias assessment for included studies, and qualitative or quantitative synthesis of results. Relevance of published reviews meeting these criteria will be assessed based on comparability of PICOTS and the extent to which included studies reflect current practice. The quality of relevant existing reviews will be graded using a components approach, with key components including search of multiple sources, use of a generally accepted tool for risk of bias assessment, and sufficient information to assess the strength of the body of evidence that includes the major domains of risk of bias, directness, consistency, precision and reporting bias. If the risk of bias assessments from the existing review are compatible with our component based approach, we will use these assessments where feasible after reviewing a sample of studies to confirm concordance with our approach—in the event the approaches are not concordant, we will perform an independent synthesis of all studies meeting our specified inclusion criteria. Key aspects of previous reviews to be described include number and types of studies included, strength of evidence assessment, and overall qualitative or quantitative findings. Newly identified studies will be presented separately from the results of existing reviews. Overall strength of evidence findings will be based on the body of evidence based on the primary evidence, not the quality or number of existing reviews.

Assessing Applicability

We will assess applicability across our key questions using the method described in AHRQ's Methods Guide.[57,65] 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 are different across studies that recruit different populations (e.g., age groups, U.S. vs. non-U.S. settings) or use 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 will use a checklist to guide the assessment of applicability. We will use these data to evaluate the applicability to clinical practice, paying special attention to study eligibility criteria, demographic features of the enrolled population in comparison to the target population, characteristics of the intervention used in comparison with care models currently in use, the possibility of surgical learning curves, and clinical relevance and timing of the outcome measures. We will summarize issues of applicability qualitatively.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

V. References

1. Gnoth C, Godehardt E, Frank-Herrmann P, et al. Definition and prevalence of subfertility and infertility. Hum Reprod. 2005;20(5):1144-7. PMID: 15802321.
2. Chandra A, Copen CE, Stephen EH. Infertility and impaired fecundity in the United States, 1982-2010: data from the National Survey of Family Growth. Natl Health Stat Report. 2013(67):1-18. PMID: 24988820.
3. Simonsen SE, Baksh L, Stanford JB. Infertility treatment in a population-based sample: 2004-2005. Matern Child Health J. 2012;16(4):877-86. PMID: 21559776.
4. Greil AL, McQuillan J, Shreffler KM, et al. Race-ethnicity and medical services for infertility: stratified reproduction in a population-based sample of U.S. women. J Health Soc Behav. 2011;52(4):493-509. PMID: 22031500.
5. Hammoud AO, Gibson M, Stanford J, et al. In vitro fertilization availability and utilization in the United States: a study of demographic, social, and economic factors. Fertil Steril. 2009;91(5):1630-5. PMID: 18539275.
6. Macaluso M, Wright-Schnapp TJ, Chandra A, et al. A public health focus on infertility prevention, detection, and management. Fertil Steril. 2010;93(1):16.e1-.e0. PMID: 18992879.
7. Sunderam S, Kissin DM, Flowers L, et al. Assisted reproductive technology surveillance—United States, 2009. MMWR Surveill Summ. 2012;61(7):1-23. PMID: 23114281.
8. Louis JF, Thoma ME, Sorensen DN, et al. The prevalence of couple infertility in the United States from a male perspective: evidence from a nationally representative sample. Andrology. 2013;1(5):741-8. PMID: 23843214.
9. Eijkemans MJ, van Poppel F, Habbema DF, et al. Too old to have children? Lessons from natural fertility populations. Hum Reprod. 2014;29(6):1304-12. PMID: 24676403.
10. Maheshwari A, Hamilton M, Bhattacharya S. Effect of female age on the diagnostic categories of infertility. Hum Reprod. 2008;23(3):538-42. PMID: 18308834.
11. Centers for Disease Control and Prevention, American Society for Reproductive Medicine, Society for Assisted Reproductive Technology. 2012 Assisted Reproductive Technology National Summary Report. Atlanta: U.S. Dept of Health and Human Services; 2014. www.cdc.gov/art/pdf/2012-report/national-summary/art_2012_national_summary_report.pdf. Accessed May 28, 2015.
12. Barton SE, Najita JS, Ginsburg ES, et al. Infertility, infertility treatment, and achievement of pregnancy in female survivors of childhood cancer: a report from the Childhood Cancer Survivor Study cohort. Lancet Oncol. 2013;14(9):873-81. PMID: 23856401.
13. Howard-Anderson J, Ganz PA, Bower JE, et al. Quality of life, fertility concerns, and behavioral health outcomes in younger breast cancer survivors: a systematic review. J Natl Cancer Inst. 2012;104(5):386-405. PMID: 22271773.
14. Letourneau JM, Ebbel EE, Katz PP, et al. Pretreatment fertility counseling and fertility preservation improve quality of life in reproductive age women with cancer. Cancer. 2012;118(6):1710-7. PMID: 21887678.
15. Odisho AY, Nangia AK, Katz PP, et al. Temporal and geospatial trends in male factor infertility with assisted reproductive technology in the United States from 1999-2010. Fertil Steril. 2014;102(2):469-75. PMID: 24931206.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

16. Practice Committee of American Society for Reproductive Medicine. Diagnostic evaluation of the infertile female: a committee opinion. Fertil Steril. 2012;98(2):302-7. PMID: 22698637.
17. National Collaborating Centre for Women's and Children's Health (UK). Fertility: Assessment and Treatment for People with Fertility Problems. London (UK): RCOG Press; 2004 Feb. (NICE Clinical Guidelines, No. 11.) Available at: http://www.ncbi.nlm.nih.gov/books/NBK45935/. Accessed April 8, 2015.
18. Hwang K, Walters RC, Lipshultz LI. Contemporary concepts in the evaluation and management of male infertility. Nat Rev Urol. 2011;8(2):86-94. PMID: 21243017.
19. Beck S, Wojdyla D, Say L, et al. The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. Bull World Health Organ. 2010;88(1):31-8. PMID: 20428351.
20. Legro RS, Myers E. Surrogate end-points or primary outcomes in clinical trials in women with polycystic ovary syndrome? Hum Reprod. 2004;19(8):1697-704. PMID: 15192061.
21. Messerlian C, Maclagan L, Basso O. Infertility and the risk of adverse pregnancy outcomes: a systematic review and meta-analysis. Hum Reprod. 2013;28(1):125-37. PMID: 23042798.
22. Barnhart KT. Live birth is the correct outcome for clinical trials evaluating therapy for the infertile couple. Fertil Steril. 2014;101(5):1205-8. PMID: 24786740.
23. Legro RS, Wu X, Barnhart KT, et al. Improving the reporting of clinical trials of infertility treatments (IMPRINT): modifying the CONSORT statement. Hum Reprod. 2014;29(10):2075-82. PMID: 25217611.
24. Kurta ML, Moysich KB, Weissfeld JL, et al. Use of fertility drugs and risk of ovarian cancer: results from a U.S.-based case-control study. Cancer Epidemiol Biomarkers Prev. 2012;21(8):1282-92. PMID: 22707710.
25. Rizzuto I, Behrens RF, Smith LA. Risk of ovarian cancer in women treated with ovarian stimulating drugs for infertility. Cochrane Database of Systematic Reviews. 2013;8:CD008215. PMID: 23943232.
26. Trabert B, Lamb EJ, Scoccia B, et al. Ovulation-inducing drugs and ovarian cancer risk: results from an extended follow-up of a large United States infertility cohort. Fertil Steril. 2013;100(6):1660-6. PMID: 24011610.
27. Kawwass JF, Crawford S, Kissin DM, et al. Tubal factor infertility and perinatal risk after assisted reproductive technology. Obstet Gynecol. 2013;121(6):1263-71. PMID: 23812461.
28. Stern JE, Luke B, Tobias M, et al. Adverse pregnancy and birth outcomes associated with underlying diagnosis with and without assisted reproductive technology treatment. Fertil Steril. 2015. PMID: 25813277.
29. Manipalviratn S, DeCherney A, Segars J. Imprinting disorders and assisted reproductive technology. Fertil Steril. 2009;91(2):305-15. PMID: 19201275.
30. Batcheller A, Cardozo E, Maguire M, et al. Are there subtle genome-wide epigenetic alterations in normal offspring conceived by assisted reproductive technologies? Fertil Steril. 2011;96(6):1306-11. PMID: 22035969.
31. Greil AL, McQuillan J, Lowry M, et al. Infertility treatment and fertility-specific distress: A longitudinal analysis of a population-based sample of U.S. women. Soc Sci Med. 2011;73(1):87-94. PMID: 21645954.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

32. Wilkins KM, Warnock JK, Serrano E. Depressive symptoms related to infertility and infertility treatments. Psychiatr Clin North Am. 2010;33(2):309-21. PMID: 20385339.
33. Feinberg EC, Larsen FW, Wah RM, et al. Economics may not explain Hispanic underutilization of assisted reproductive technology services. Fertil Steril. 2007;88(5):1439-41. PMID: 17561005.
34. Fujimoto VY, Luke B, Brown MB, et al. Racial and ethnic disparities in assisted reproductive technology outcomes in the United States. Fertil Steril. 2010;93(2):382-90. PMID: 19081561.
35. McCarthy-Keith DM, Schisterman EF, Robinson RD, et al. Will decreasing assisted reproduction technology costs improve utilization and outcomes among minority women? Fertil Steril. 2010;94(7):2587-9. PMID: 20356585.
36. Wellons MF, Fujimoto VY, Baker VL, et al. Race matters: a systematic review of racial/ethnic disparity in Society for Assisted Reproductive Technology reported outcomes. Fertil Steril. 2012;98(2):406-9. PMID: 22698638.
37. Kawwass JF, Monsour M, Crawford S, et al. Trends and outcomes for donor oocyte cycles in the United States, 2000-2010. JAMA. 2013;310(22):2426-34. PMID: 24135860.
38. Myers ER. Outcomes of donor oocyte cycles in assisted reproduction. JAMA. 2013;310(22):2403-4. PMID: 24135802.
39. Klitzman RL, Sauer MV. Kamakahi vs ASRM and the future of compensation for human eggs. Am J Obstet Gynecol. 2015. PMID: 25816784.
40. De Melo-Martin I. The ethics of anonymous gamete donation: is there a right to know one's genetic origins? Hastings Cent Rep. 2014;44(2):28-35. PMID: 24532424.
41. Wu AK, Odisho AY, Washington SL, 3rd, et al. Out-of-pocket fertility patient expense: data from a multicenter prospective infertility cohort. J Urol. 2014;191(2):427-32. PMID: 24018235.
42. Bitler M, Schmidt L. Health disparities and infertility: impacts of state-level insurance mandates. Fertil Steril. 2006;85(4):858-65. PMID: 16580365.
43. Bitler MP, Schmidt L. Utilization of infertility treatments: the effects of insurance mandates. Demography. 2012;49(1):125-49. PMID: 22167581.
44. Hamilton BH, McManus B. The effects of insurance mandates on choices and outcomes in infertility treatment markets. Health Econ. 2012;21(8):994-1016. PMID: 21905150.
45. Kiatpongsan S, Huckman RS, Hornstein MD. The Great Recession, insurance mandates, and the use of in vitro fertilization services in the United States. Fertil Steril. 2015;103(2):448-54. PMID: 25497450.
46. Practice Committee of Society for Assisted Reproductive Technology, Practice Committee of American Society for Reproductive Medicine. Elective single-embryo transfer. Fertil Steril. 2012;97(4):835-42. PMID: 22196716.
47. Garrido N, Bellver J, Remohi J, et al. Cumulative newborn rates increase with the total number of transferred embryos according to an analysis of 15,792 ovum donation cycles. Fertil Steril. 2012;98(2):341-6.e1-2. PMID: 22633263.
48. Dar S, Librach CL, Gunby J, et al. Increased risk of preterm birth in singleton pregnancies after blastocyst versus Day 3 embryo transfer: Canadian ART Register (CARTR) analysis. Hum Reprod. 2013;28(4):924-8. PMID: 23349411.
49. Barnhart KT. Introduction: are we ready to eliminate the transfer of fresh embryos in in vitro fertilization? Fertil Steril. 2014;102(1):1-2. PMID: 24890272.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

50. Roque M, Lattes K, Serra S, et al. Fresh embryo transfer versus frozen embryo transfer in in vitro fertilization cycles: a systematic review and meta-analysis. Fertil Steril. 2013;99(1):156-62. PMID: 23040524.
51. Kulkarni AD, Jamieson DJ, Jones HW, Jr., et al. Fertility treatments and multiple births in the United States. N Engl J Med. 2013;369(23):2218-25. PMID: 24304051.
52. Reindollar RH, Regan MM, Neumann PJ, et al. A randomized clinical trial to evaluate optimal treatment for unexplained infertility: the fast track and standard treatment (FASTT) trial. Fertil Steril. 2010;94(3):888-99. PMID: 19531445.
53. Huang LN, Tan J, Hitkari J, et al. Should IVF be used as first-line treatment or as a last resort? A debate presented at the 2013 Canadian Fertility and Andrology Society meeting. Reprod Biomed Online. 2015;30(2):128-36. PMID: 25498596.
54. Kawwass JF, Kissin DM, Kulkarni AD, et al. Safety of assisted reproductive technology in the United States, 2000-2011. JAMA. 2015;313(1):88-90. PMID: 25562272.
55. Myers ER, McCrory DC, Mills AA, et al. Effectiveness of Assisted Reproductive Technology. Evidence Report/Technology Assessment No. 167 (Prepared by the Duke University Evidence-based Practice Center under Contract No. 290-02-0025.) AHRQ Publication No. 08-E012. Rockville, MD: Agency for Healthcare Research and Quality. May 2008. Available at: www.ahrq.gov/research/findings/evidence-based-reports/infertil-evidence-report.pdf. Accessed April 8, 2015.
56. Myers ER, McCrory DC, Mills AA, et al. Effectiveness of assisted reproductive technology (ART). Evid Rep Technol Assess (Full Rep). 2008(167):1-195. PMID: 18620469.
57. Agency for Healthcare Research and Quality. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville, MD: Agency for Healthcare Research and Quality. Available at: http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318. Accessed July 16, 2015.
58. Myers ER, McCrory DC, Mills AA, et al. Effectiveness of Assisted Reproductive Technology. Evidence Report/Technology Assessment No. 167 (Prepared by the Duke University Evidence-based Practice Center under Contract No. 290-02-0025.) AHRQ Publication No. 08-E012. Rockville, MD: Agency for Healthcare Research and Quality. May 2008. Available at: http://www.ahrq.gov/research/findings/evidence-based-reports/infertil-evidence-report.pdf. Accessed July 17, 2015.
59. Robinson KA, Chou R, Berkman ND, et al. Integrating Bodies of Evidence: Existing Systematic Reviews and Primary Studies. Methods Guide for Comparative Effectiveness Reviews (Prepared by the Scientific Resource Center under Contract No. 290-2012-00004-C). AHRQ Publication No. 15-EHC007-EF. Rockville, MD: Agency for Healthcare Research and Quality. February 2015. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.
60. Myers E, Sanders GD, Ravi D, et al. Evaluating the Potential Use of Modeling and Value-of-Information Analysis for Future Research Prioritization Within the Evidence-based Practice Center Program. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 11-EHC030-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2011. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

61. Myers ER, Misurski DA, Swamy GK. Influence of timing of seasonal influenza vaccination on effectiveness and cost-effectiveness in pregnancy. Am J Obstet Gynecol. 2011;204(6 Suppl 1):S128-40. PMID: 21640230.
62. Brown University Evidence-based Practice Center. Guidance for the Conduct and Reporting of Decision and Simulation Models in the Context of Health Technology Assessment. Research Report - Draft – Sept. 25, 2014 [full citation pending]. Available at: http://effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?productid=1979&pageaction=displayproduct.
63. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions-Agency for Healthcare Research and Quality and the Effective Health Care Program. J Clin Epidemiol. 2010;63(5):513-23. PMID: 19595577.
64. Myers ER, Moorman P, Gierisch JM, et al. Benefits and harms of breast cancer screening: a systematic review. JAMA (In press).
65. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011;64(11):1198-207. PMID: 21463926.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

VI. Definition of Terms

ART	assisted reproductive technology
BMI	body mass index
GI	gastrointestinal
GnRH	gonadotropin-releasing hormone
GRADE	Grading of Recommendations Assessment, Development and Evaluation
hCG	human chorionic gonadotropin
ICSI	intra-cytoplasmic sperm injection
IUI	intrauterine insemination
IVF	*in vitro* fertilization
KQ(s)	key question(s)
NASS	National ART Surveillance System
OHSS	ovarian hyperstimulation syndrome
PCOS	polycystic ovary syndrome
PICOTS	population, interventions, comparators, outcomes, timings, and settings
RCT(s)	randomized controlled trials(s)

VII. Summary of Protocol Amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol.

VIII. Review of Key Questions

AHRQ posted the key questions on the Effective Health Care Website for public comment. The EPC refined and finalized the key questions after review of the public comments, and input from Key Informants and the Technical Expert Panel (TEP). This input is intended to ensure that the key questions are specific and relevant.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Experts

Technical Experts constitute a multi-disciplinary group of clinical, content, and methodological experts who provide input in defining populations, interventions, comparisons, or outcomes and identify particular studies or databases to search. They are selected to provide broad expertise and perspectives specific to the topic under development. Divergent and conflicting opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design, and methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as requested by the EPC. Technical Experts do not do analysis of any kind nor do they contribute to the writing of the report. They have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodological expertise. The EPC considers all peer review comments on the draft report in preparation of the final report. Peer reviewers do not participate in writing or editing of the final report or other products. The final report does not necessarily represent the views of individual reviewers. The EPC will complete a disposition of all peer review comments. The disposition of comments for systematic reviews and technical briefs will be published three months after the publication of the evidence report.

Potential Peer Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest that cumulatively total greater than $1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder

This project was funded under Contract No. xxx-xxx from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Source: www.effectivehealthcare.ahrq.gov
Published online: October 12, 2015