Methods Research Report

Supplemental Project To Assess the Transparency of Reporting for Trials Evaluating Treatment for Infertility

Methods Research Report

Supplemental Project To Assess the Transparency of Reporting for Trials Evaluating Treatment for 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:
John W. Williams Jr., M.D., M.H.Sc.
Jennifer L. Eaton, M.D.
Jennifer M. Gierisch, Ph.D.
Varsha Masilamani, M.B.B.S.
Megan von Isenburg, M.S.L.S.
Megan Chobot, M.S.L.S.

AHRQ Publication No. 17-EHC022-EF
June 2017

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 health care decision makers—patients and clinicians, health system leaders, and policymakers, among others—make well-informed decisions and thereby improve the quality of health care 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.

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

Suggested citation: Williams JW Jr., Eaton JL, Gierisch JM, Masilamani V, von Isenburg M, Chobot M. Supplemental Project To Assess the Transparency of Reporting for Trials Evaluating Treatment for Infertility. Methods Research Report. (Prepared by the Duke University Evidence-based Practice Center under Contract No. 290-2015-00004-I) AHRQ Publication No. 17-EHC022-EF. Rockville, MD: Agency for Healthcare Research and Quality. June 2017.www.effectivehealthcare.ahrq.gov/reports/final.cfm. DOI: https://doi.org/10.23970/AHRQEPCMETH4.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies and strategies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To improve the scientific rigor of these evidence reports, AHRQ supports empiric research by the EPCs to help understand or improve complex methodologic issues in systematic reviews. These methods research projects are intended to contribute to the research base in and be used to improve the science of systematic reviews. They are not intended to be guidance to the EPC program, although may be considered by EPCs along with other scientific research when determining EPC program methods guidance.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality. The reports undergo peer review prior to their release as a final report.

If you have comments on this Methods Research Project 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 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

Elise Berliner, Ph.D.
Task Order Officer
Center for Evidence and Practice Improvement
Agency for Healthcare Research and Quality

Acknowledgments

The authors thank Gillian Sanders-Schmidler, Ph.D., and Evan Myers, M.D., M.P.H, for technical input and coordination with the associated Management of Infertility systematic review.

Contents

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 is 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 systematic review (SR) the Duke EPC is preparing, entitled Management of Infertility, to explore differences between information from published sources included in the review and CT.gov. The objectives of this SR are to evaluate the comparative safety and effectiveness of treatment strategies for: a) women of reproductive age (18-44) who are subfertile/infertile due to polycystic ovary syndrome (PCOS), endometriosis, unknown reasons, or tubal or peritoneal factors; or b) couples with male factor infertility; and evaluate short- and long-term health outcomes of gamete donors in infertility.

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 Key Question (KQ) 1, KQ 2, and KQ 4 from the Management of Infertility SR 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 SR 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 of the Management of Infertility SR.

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 SR at several levels.

First, we determined whether RCTs reporting a live birth outcome that were included in the Management of Infertility SR 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 the bibliographical database (EndNote® Version X7; Thomson Reuters, Philadelphia, PA). 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 SR 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 SR
- 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 SR. 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 SR, 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 SR and in CT.gov; (2) studies identified from CT.gov as potentially eligible but not included in the Management of Infertility SR; (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 SR 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 SR. 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 SR

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 SR, 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 SR search and included in the review. The other five studies were reviewed by an investigator; details are reported in the Table 1.

Table 1. Potentially eligible studies not included in the review

NCTID	Search Strategy Identifying Trial	CT.gov Completion Date	Classification
NCT01675843	Both	March 2012	Potentially eligible; no citation in PubMed
NCT01679574	Both	January 2012	Potentially eligible; no citation in PubMed
NCT01894074	Broad	July 2015	Potentially eligible, no citation in PubMed
NCT00220545	Both	March 2006	Identified in original review search but excluded at title-and-abstract

NCTID	Search Strategy Identifying Trial	CT.gov Completion Date	Classification
			screening stage. Full text reviewed and study included in Management of Infertility SR.
NCT01581359	Both	May 2015	Potentially eligible, no citation in PubMed

Only five potentially eligible studies were identified across the three KQs. Of these, two 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 eight studies matched by the NCTID. These data were compared to data abstracted from published data by the Management of Infertility SR 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 eight studies.
- The study design and number of study arms matched for all eight studies.
- Of five studies reporting the enrolled "n," four were exact matches and one had a discrepancy in the estimated enrollment (326) versus 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 eight 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 were classified as matched for six studies. Two studies were classified as discrepant: one of these was classified as non-government/non-industry from CT.gov and as "not reported" from manuscript, and one was classified as non-government/non-industry from CT.gov and as government from the published manuscript.

Outcomes were compared at two 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,[7] miscarriage,[8] and live birth.[9] In four studies, outcomes reported in published manuscripts were not described in CT.gov: live birth,[8,10] miscarriage,[11] multiple births,[11,12] and surgical complications.[12]

- Only one[7] of the eight 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 two 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-one trials (5,718 patients), including one study reclassified as eligible (see Table 1), were included in the SOE rating for KQ 1, and thus these two "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 three 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 versus staff are given in Table 2.

Table 2. Person-hours required, investigators versus staff

Name	Administrative (meetings, etc.)	Planning/ designing	Running searches/ abstracting data	Synthesizing data/writing	Total
Investigator	7	9	10.75	7.75	34.5
EPC Staff	23	0	14	3	40
Totals	30	9	24.75	10.75	74.5

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 SR were matched to a CT.gov record based on NCTID, and only one 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 two studies without a publication and one 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

1. Borenstein M, Hedges LV, Higgins JPT, et al. Publication Bias. Introduction to Meta-Analysis. John Wiley & Sons, Ltd; 2009:277-92.
2. Song F, Parekh S, Hooper L, et al. Dissemination and publication of research findings: an updated review of related biases. Health Technology Assessment. 2010 Feb;14(8):iii, ix-xi, 1-193. doi: 10.3310/hta14080. PMID: 20181324.
3. Sterne JA, Sutton AJ, Ioannidis JP, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ. 2011;343:d4002. doi: 10.1136/bmj.d4002. PMID: 21784880.
4. Dwan K, Altman DG, Clarke M, et al. Evidence for the selective reporting of analyses and discrepancies in clinical trials: a systematic review of cohort studies of clinical trials. PLoS Med. 2014 Jun;11(6):e1001666. doi: 10.1371/journal.pmed.1001666. PMID: 24959719.
5. Hartung DM, Zarin DA, Guise JM, et al. Reporting discrepancies between the ClinicalTrials.gov results database and peer-reviewed publications. Ann Intern Med. 2014 Apr 1;160(7):477-83. doi: 10.7326/M13-0480. PMID: 24687070.
6. Agency for Healthcare Research and Quality (AHRQ). Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Available at: http://www.effectivehealthcare.ahrq.gov/methodsguide.cfm. Accessed April 1, 2016.
7. Legro RS, Brzyski RG, Diamond MP, et al. Letrozole versus clomiphene for infertility in the polycystic ovary syndrome. N Engl J Med. 2014 Jul 10;371(2):119-29. doi: 10.1056/NEJMoa1313517. PMID: 25006718.
8. Morin-Papunen L, Rantala AS, Unkila-Kallio L, et al. Metformin improves pregnancy and live-birth rates in women with polycystic ovary syndrome (PCOS): a multicenter, double-blind, placebo-controlled randomized trial. J Clin Endocrinol Metab. 2012 May;97(5):1492-500. doi: 10.1210/jc.2011-3061. PMID: 22419702.
9. Palomba S, Falbo A, Carrillo L, et al. Metformin reduces risk of ovarian hyperstimulation syndrome in patients with polycystic ovary syndrome during gonadotropin-stimulated in vitro fertilization cycles: a randomized, controlled trial. Fertil Steril. 2011 Dec;96(6):1384-90.e4. doi: 10.1016/j.fertnstert.2011.09.020. PMID: 21982727.
10. Abu Hashim H, El Rakhawy M, Abd Elaal I. Randomized comparison of superovulation with letrozole vs. clomiphene citrate in an IUI program for women with recently surgically treated minimal to mild endometriosis. Acta Obstet Gynecol Scand. 2012 Mar;91(3):338-45. doi: 10.1111/j.1600-0412.2011.01346.x. PMID: 22181973.
11. Johnson NP, Stewart AW, Falkiner J, et al. PCOSMIC: a multi-centre randomized trial in women with PolyCystic Ovary Syndrome evaluating Metformin for Infertility with Clomiphene. Hum Reprod. 2010 Jul;25(7):1675-83. doi: 10.1093/humrep/deq100. PMID: 20435692.
12. Palomba S, Falbo A, Battista L, et al. Laparoscopic ovarian diathermy vs clomiphene citrate plus metformin as second-line strategy for infertile anovulatory patients with polycystic ovary syndrome: a randomized controlled trial. Am J Obstet Gynecol. 2010 Jun;202(6):577.e1-8. doi: 10.1016/j.ajog.2009.11.042. PMID: 20096821.