







# A protocol developing, disseminating and implementing a core outcome set for infertility

J.M.N. Duffy <sup>1,2,\*</sup>, S. Bhattacharya<sup>3</sup>, C. Curtis<sup>4,5</sup>, J.L.H. Evers <sup>6</sup>, R.G. Farquharson<sup>7</sup>, S. Franik<sup>8</sup>, Y. Khalaf<sup>9</sup>, R.S. Legro<sup>10</sup>, S. Lensen<sup>11</sup>, B.W. Mol<sup>12</sup>, C. Niederberger<sup>13</sup>, E.H.Y. Ng<sup>14</sup>, S. Repping<sup>15</sup>, A. Strandell<sup>16</sup>, H.L. Torrance<sup>17</sup>, A. Vail<sup>18</sup>, M. van Wely<sup>15</sup>, N.L. Vuong <sup>19</sup>, A.Y. Wang <sup>20</sup>, R. Wang <sup>21</sup>, J. Wilkinson <sup>18</sup>, M.A. Youssef<sup>22</sup>, and C.M. Farquhar<sup>11</sup>, on behalf of COMMIT: Core Outcomes Measures for Infertility Trials

<sup>1</sup>Balliol College, University of Oxford, Oxford, UK <sup>2</sup>Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK <sup>3</sup>Institute of Applied Health Sciences, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK <sup>4</sup>Fertility New Zealand, Auckland, New Zealand <sup>5</sup>School of Psychology, University of Waikato, Hamilton, New Zealand <sup>6</sup>Centre for Reproductive Medicine and Biology, University Medical Centre Maastricht, Maastricht, The Netherlands <sup>7</sup>Department of Obstetrics and Gynaecology, Liverpool Women's NHS Foundation Trust, Liverpool, UK <sup>8</sup>Department of Obstetrics and Gynaecology, Münster University Hospital, Münster, Germany <sup>9</sup>Assisted Conception Unit, Guy's Hospital, London, UK <sup>10</sup>Department of Obstetrics and Gynaecology, Penn State College of Medicine, PA, USA <sup>11</sup>Department of Obstetrics and Gynaecology, University of Auckland, Auckland, New Zealand <sup>12</sup>Department of Obstetrics and Gynaecology, School of Medicine, Monash University, Melbourne, Australia <sup>13</sup>Department of Urology, University of Illinois at Chicago College of Medicine, Chicago, IL, USA <sup>14</sup>Department of Obstetrics and Gynaecology, The University of Hong Kong, Hong Kong <sup>15</sup>Center for Reproductive Medicine, Amsterdam Reproduction and Development Institute, Academic Medical Centre, Amsterdam, The Netherlands <sup>16</sup>Sahlgrenska University Hospital, Göteborg, Sweden <sup>17</sup>Department of Reproductive Medicine, University Medical Centre Utrecht, Utrecht, The Netherlands <sup>18</sup>Centre for Biostatistics, University of Manchester, Manchester, UK <sup>19</sup>Department of Obstetrics and Gynaecology, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam <sup>20</sup>Faculty of Health, University of Technology Sydney, Broadway, Australia <sup>21</sup>Robinson Research Institute and Adelaide Medical School, University of Adelaide, Adelaide, Australia <sup>22</sup>Department of Obstetrics & Gynaecology, Faculty of Medicine, Cairo University, Cairo, Egypt

\*Correspondence address. Balliol College, University of Oxford, Oxford OX1 3BJ, UK. E-mail: james.duffy@balliol.ox.ac.uk  [orcid.org/0000-0001-6127-9859](https://orcid.org/0000-0001-6127-9859)

Submitted on January 20, 2018; editorial decision on March 14, 2018; accepted on April 12, 2018

**STUDY QUESTIONS:** We aim to produce, disseminate and implement a core outcome set for future infertility research.

**WHAT IS KNOWN ALREADY:** Randomized controlled trials (RCTs) evaluating infertility treatments have reported many different outcomes, which are often defined and measured in different ways. Such variation contributes to an inability to compare, contrast and combine results of individual RCTs. The development of a core outcome set will ensure outcomes important to key stakeholders are consistently collected and reported across future infertility research.

**STUDY DESIGN, SIZE, DURATION:** This is a consensus study using the modified Delphi method. All stakeholders, including healthcare professionals, allied healthcare professionals, researchers and people with lived experience of infertility will be invited to participate.

**PARTICIPANTS/MATERIALS, SETTING, METHODS:** An international steering group, including people with lived experience of infertility, healthcare professionals, allied healthcare professionals and researchers, has been formed to guide the development of this core outcome set. Potential core outcomes have been identified through a comprehensive literature review of RCTs evaluating treatments for infertility and will be entered into a modified Delphi method. Participants will be asked to score potential core outcomes on a nine-point Likert scale anchored between one (not important) and nine (critical). Repeated reflection and rescaling should promote convergence towards consensus 'core' outcomes. We will establish standardized definitions and recommend high-quality measurement instruments for individual core outcomes.

**STUDY FUNDING/COMPETING INTEREST(S):** This project is funded by the Royal Society of New Zealand Catalyst Fund (3712235). BWM reports consultancy fees from Guerbet, Merck, and ObsEva. R.S.L. reports consultancy fees from Abbvie, Bayer, Fractyl and Ogeda and research sponsorship from Ferring. S.B. is the Editor-in-Chief of Human Reproduction Open. The remaining authors declare no competing interests.

**Key words:** infertility / consensus study / modified Delphi method / core outcome set / randomized controlled trials / systematic review / clinical practice guidelines

## WHAT DOES THIS MEAN FOR PATIENTS?

Research studies testing new treatments for infertility often measure different outcomes.

Some use pregnancy as their mark of success, while others count live births, and they may, or may not, collect information about risks or side effects. When complete, the results from different research studies cannot be easily compared or combined, to see which treatments work best. This is a barrier to improving the care people with infertility receive.

The article explains that an international group of healthcare professionals, researchers and fertility patients has been set up to overcome this barrier by developing a core set of outcomes that would be common to all future infertility research. They started by collecting all the different outcomes currently reported by infertility research and will go on to ask healthcare professionals, allied healthcare professionals, researchers and fertility patients what outcomes they think are important in a three-round questionnaire. Anyone, anywhere can participate. The results will be used to select the most meaningful outcomes which will consistently be collected and reported in future infertility research.

## Introduction

Infertility is a disease of the reproductive system defined by the failure to achieve a clinical pregnancy following 12 months or more of regular unprotected sexual intercourse (Zegers-Hochschild et al., 2009). It is estimated that infertility affects one in six couples (Thoma et al., 2013). The main categories of infertility include ovulatory disorders, tubal damage, uterine or peritoneal disorders, reduced semen quantity and quality, and unexplained infertility (Collins and van Steirteghem, 2004). Treatment falls into two main categories: interventions to restore fertility when a clear cause is established, such as ovulation induction, and assisted reproductive technology (ART). Potential infertility treatments require careful evaluation.

Randomized controlled trials (RCTs), where the inclusion criteria are broad, and the outcomes are patient-centred, are the best way of establishing the comparative efficacy and safety of treatments. While individual RCTs are useful, pooling results from all available RCTs is likely to provide the best evidence to inform clinical practice (Duffy et al., 2017a). Considerable attention has been paid to standardizing the methods of conducting RCTs and systematic reviews. However, the selection, collection and reporting of outcomes has been largely overlooked and there is currently limited consensus regarding the outcomes and outcome measures infertility research should collect and report (Duffy et al., 2017c).

In the absence of a standardized approach, researchers have made arbitrary decisions, and this has resulted in many different outcomes and outcome measures collected and reported in infertility RCTs (Dapuzzo et al., 2011; Braakhekke et al., 2014; Wilkinson et al., 2016). A recent systematic review of 910 RCTs evaluating fertility treatments identified that only a fifth of included trials reported live birth (182/910; 20%). Singleton live birth was the primary outcome in 68 trials (7.4%). Only a minority of included trials reported maternal outcomes (52/910; 5.7%) and neonatal outcomes (44/910; 4.8%) (Braakhekke

et al., 2014). Adverse outcomes, including ovarian hyperstimulation syndrome, congenital anomalies and long-term health risks, were infrequently reported (Braakhekke et al., 2015). The complexity of the situation is further increased because individual numerators can be associated with numerous different denominators: for example, a recent review of infertility RCTs identified 15 different denominators for clinical pregnancy (Wilkinson et al., 2017).

The lack of consensus regarding the collection and reporting of outcomes limits the comparison and pooling of individual trial data and directly impacts the usefulness of research to inform patients and professionals in clinical practice. Such diversity of outcome collection and reporting has been demonstrated in many different reproductive conditions, including endometriosis, pre-eclampsia and preterm birth (Hirsch et al., 2016; Duffy et al., 2017b, 2017c, 2017d; van't Hooff et al., 2016).

With so many different outcomes and outcome measures in common use, there is considerable scope for selective emphasis and reporting of results. Outcome reporting bias is defined as the selection for publication of a subset of the original recorded outcome variables on the basis of the results (Kirkham et al., 2010). Several systematic reviews evaluating interventions for infertility have reported the possibility of outcome reporting bias (Duffy et al., 2009, 2010, 2014). Several studies have confirmed outcome reporting bias and quantified its impact when pooling data from individual RCTs in a meta-analysis (Chan et al., 2004; Chan and Altman, 2005; Kirkham et al., 2010; Smyth et al., 2011; Hart et al., 2012; Page et al., 2014; Saini et al., 2014). A systematic review of 157 Cochrane systematic reviews, published in 2007, revealed that over a third had at least one randomized trial at high risk of outcome reporting bias (Kirkham et al., 2010). A sensitivity analysis, excluding those with reporting bias, demonstrated a relative reduction of over 20% in the treatment effect of the primary outcome (Clark et al., 1998).

The research community has engaged with the process of standardizing the reporting of RCTs of infertility treatments. An international

working group established a consensus process to modify the Consolidated Standards of Reporting Trials (CONSORT) statement. The improving the reporting of RCTs of infertility treatments (IMPRINT) statement made several recommendations including a preferred primary outcome for all infertility RCTs of live birth, defined as delivery of a live infant >20 weeks gestation, or cumulative live birth, defined as where more than one cycle occurs or where frozen embryos are transferred (Legro *et al.*, 2014).

The next challenge is to address the unwarranted variation in outcome collection and reporting (Dapuzzo *et al.*, 2011; Braakhekke *et al.*, 2014; Wilkinson *et al.*, 2016). The development and implementation of a minimum data set, termed a core outcome set, would help to address these issues. Core outcome sets are minimum collections of outcomes with standardized measurement and reporting (Williamson *et al.*, 2012). They are identified using consensus science methods, thereby enabling key stakeholders, including healthcare professionals, researchers and patients, to suggest and prioritize outcomes (Duffy *et al.*, 2017e).

We aim to produce, disseminate and implement a core outcome set for infertility RCTs, systematic reviews and clinical practice guidelines.

## Materials and Methods

### Prospective registration

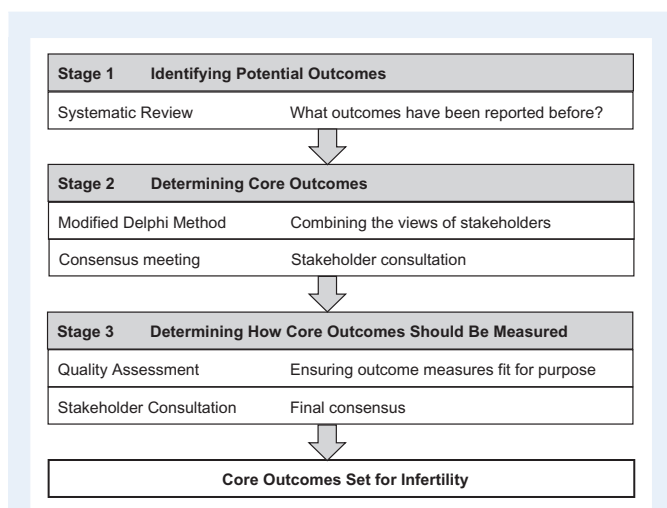
This study has been prospectively registered with the Core Outcome Measures in Effectiveness Trials (COMET) initiative; the registration number is 1023 and is available online ([www.comet-initiative.org/studies/details/1023](http://www.comet-initiative.org/studies/details/1023)).

### Steering group

An international steering group, including people with lived experience of infertility, healthcare professionals and researchers, has been formed to guide the development of this core outcome set (Fig. 1). Members of the steering group represent various disciplines, geographical areas and expertise.

### Core outcome set scope

The steering group recommend the core outcome set should apply to RCTs, systematic reviews and clinical practice guidelines evaluating interventions to restore or circumvent infertility.



**Figure 1** Developing a core outcome set for infertility trials.

The study's methods have been informed by reviewing previous core outcome sets in women's and newborn health (Duffy *et al.*, 2016a; Hirsch *et al.*, 2016b; Khalil *et al.*, 2017; Webbe *et al.*, 2017; Whitehouse *et al.*, 2017).

### Stage 1: Identifying potential core outcomes

We have performed a systematic review of infertility RCTs and extracted the reported outcomes and outcome measures (Wilkinson *et al.*, 2016). A comprehensive inventory of outcomes will be developed in consultation with key stakeholders including healthcare professionals, researchers and people with lived experience of infertility (Duffy and McManus, 2016b). During the iterative development of the inventory of outcomes, consideration will be given to efficacy and safety outcomes, the aim of the intervention, for example, interventions aiming to circumvent infertility, and the timing of the intervention, for example, preconception interventions. Lay definitions will be developed for individual outcomes. The outcome inventory and lay definitions will be entered into a modified Delphi method.

### Stage 2: Identifying core outcomes

The modified Delphi method assesses the extent of agreement and then resolves disagreement (Sinha *et al.*, 2011). The survey will be piloted to ensure the ease of completion. Stakeholders including healthcare professionals, researchers and people with lived experience of infertility, will be invited to participate. There is no robust method for calculating the required sample size. We aim to recruit a minimum of 16 participants for each stakeholder group.

During round one, participants will provide their demographic details and will be allocated a unique identifier, which facilitates future anonymity. Potential core outcomes will be presented within each domain and participants will be asked to score individual outcomes using a nine-point scale from one (not important for decision making) to nine (critical for decision making) (Guyatt *et al.*, 2011). Participants will be invited to suggest additional outcomes before completing the first-round survey.

All outcomes will be carried forward from round one into round two. For each outcome, the percentage of participants scoring individual outcomes during round one at each possible response from one to nine will be calculated and tabulated for their own stakeholder group. Additional outcomes will be considered by the steering group, prompting the reformulation of round one outcomes and the inclusion of additional outcomes in round two.

During the round two survey, participants will receive their own scores and stakeholder group feedback for each round one outcome. Participants will be invited to reflect upon summarized stakeholder group feedback and their own score before rescore round one outcomes as well as scoring additional outcomes suggested by participants in round one.

All outcomes will be carried forward from round two into round three. For each outcome, the percentage of participants scoring individual outcomes during round two at each possible response from one to nine will be calculated and tabulated for each individual stakeholder group. During the round three survey, participants will receive their own scores and stakeholder group feedback for each round two outcome. Participants will be invited to reflect upon summarized stakeholder group feedback and their own score before rescore round two outcomes.

Following round three, a standardized definition will be applied to the results to identify prioritized outcomes, defined by a median score of eight in each stakeholder group. A consensus development workshop will review the round three results. The objective of the consensus workshop will be to develop a final core outcome set for infertility.

## Stage 3: Identifying and standardizing core outcome measures

Once core outcomes are agreed upon, we will determine how the outcomes should be defined and measured. Potential definitions will be inventoried across formal definition development initiatives, national and international guidelines, Cochrane systematic reviews and RCTs (Fig. 2). Potential definitions will be entered into a consensus development workshop including healthcare professionals, researchers and people with lived experience of infertility. The objective of the consensus workshop will be to identify definitions for individual core outcomes. Careful attention will be paid to the appropriate selection of both numerators and denominators. Potential measurement instruments will be inventoried across national and international guidelines, Cochrane systematic reviews and RCTs. Potential measurement instruments will be quality assessed using the COMET initiative and the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) initiative quality assessment framework (Prinsen et al., 2016). High-quality measurement instruments will be associated with core outcomes.

### Ethical review

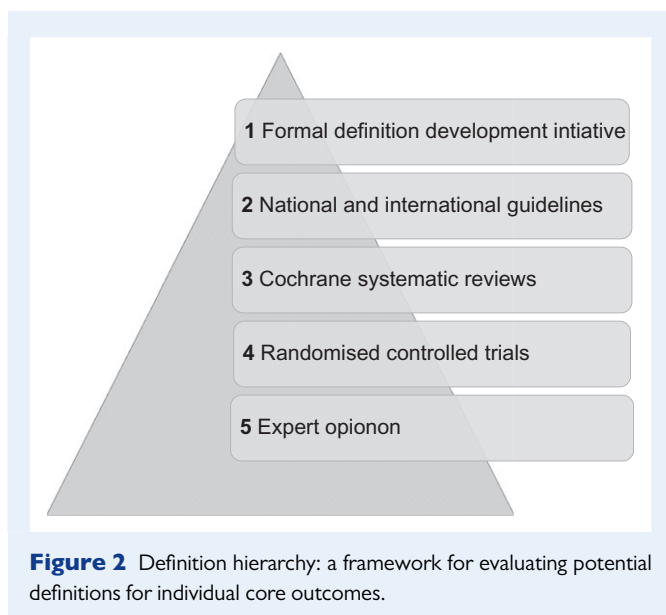
We asked the advice of the National Research Ethics Service about whether this study required ethical review by an NHS Research Ethics Committee, and they advised that this study should be considered as service evaluation and development, and therefore not require ethical review.

## Discussion

Implementing core outcome sets in future RCTs, systematic reviews and clinical guidelines could reduce research waste and advance the relevance of research to inform clinical practice.

### Improving outcome selection

The Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement, implemented by funders of health research including the National Institutes of Health, European Commission, and the National Institute of Health Research, recommend the use of



core outcome sets where they exist. A core outcome set ensures that outcomes relevant to all stakeholders, including healthcare professionals, researchers and patients, are collected and reported (Duffy et al., 2017e).

### Improving outcome reporting

The Core Outcomes in Women's and Newborn Health (CROWN) initiative, a consortium of 84 speciality journals, including the Cochrane Gynaecology and Fertility Group, Human Reproduction and Human Reproduction Update has been established to support the development, dissemination, and implementation of core outcome sets (Khan, 2014; Duffy et al., 2017e). Participating journals will require researchers to report core outcomes and outcome measures (Khan, 2014).

### Infrastructure to support future research

Developing a core outcome set will establish an international network of organizations and stakeholders with experience of contributing to a collaborative consensus study. This infrastructure could be leveraged in other settings, for example, prioritizing research uncertainties.

## Conclusion

Rigorous implementation of a core outcome set should ensure that outcomes important to all stakeholders, including healthcare professionals, researchers and people with infertility, will be collected and reported in a standardized fashion, advancing the usefulness of research to inform clinical practice.

## Acknowledgements

We would like to thank colleagues at the Cochrane Gynaecology and Fertility Group, University of Auckland, New Zealand and Mr David J. Mills for administrative and material support.

## Authors' roles

Study concept and design: J.M.D. and C.M.F. Acquisition of data: J.M.D., S.B., C.C., J.L.E., R.G.F., S.X.F., Y.K., R.E.S.L., S.L., B.W.M., C.N., N.H.Y., S.R., A.S., T.H.L., A.V., M.v.W., N.L.V., A.Y.W., R.W., M.A.Y. and C.M.F. Analysis and interpretation of data: J.M.D., S.B., C.C., J.L.E., R.G.F., S.X.F., R.E.S.L., S.L., B.W.M., C.N., N.H.Y., S.R., A.S., T.H.L., A.V., M.v.W., N.L.V., A.Y.W., R.W., M.A.Y. and C.M.F. Drafting of the article: J.M.D. and C.M.F. Critical revision of the article for important intellectual content: S.B., C.C., J.L.E., R.G.F., S.X.F., Y.K., R.E.S.L., S.L., B.W.M., C.N., N.H.Y., S.R., A.S., T.H.L., A.V., M.v.W., N.L.V., A.Y.W., R.W. and M.A.Y. Statistical analysis: A.V., R.W. and J.W. Study supervision: C.M.F.

## Funding

This project is partly funded by the Royal Society of New Zealand Catalyst Fund (3712235). The funder has no role in the design and conduct of the study, the collection, management, analysis, or interpretation of data, or article preparation. B.W.M. is supported by a National Health and Medical Research Council (Australia) Practitioner Fellowship (GNT1082548)

## Conflict of interest

B.W.M. reports consultancy fees from Guerbet, Merck and ObsEva. R.S.L. reports consultancy fees from Abbvie, Bayer, Fractyl and Ogeda and research sponsorship from Ferring. S.B. is Editor-in-Chief of Human Reproduction Open. The remaining authors declare no competing interests.

## References

- Braakhekke M, Kamphuis EI, Mol R, Norman RJ, Bhattacharya S, van der Veen F, Mol BW. Effectiveness and safety as outcome measures in reproductive medicine. *Hum Reprod* 2015;**30**:2249–2251.
- Braakhekke M, Kamphuis EI, van Rumste MM, Mol F, van der Veen F, Mol BW. How are neonatal and maternal outcomes reported in randomised controlled trials (RCTs) in reproductive medicine? *Hum Reprod* 2014;**29**: 1211–1217.
- Chan AW, Altman DG. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *Br Med J* 2005;**330**:753.
- Chan AW, Krleža-Jerić K, Schmid I, Altman DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;**171**:735–740.
- Clark DE, Smith SK, He Y, Day KA, Licence DR, Corps AN, Lammoglia R, Charnock-Jones DS. A vascular endothelial growth factor antagonist is produced by the human placenta and released into the maternal circulation. *Biol Reprod* 1998;**59**:1540–1548.
- Collins JA, Van Steirteghem A. Overall prognosis with current treatment of infertility. *Hum Reprod Update* 2004;**10**:309–316.
- Dapuzzo L, Seitz FE, Dodson WC, Stetter C, Kunselman AR, Legro RS. Incomplete and inconsistent reporting of maternal and fetal outcomes in infertility treatment trials. *Fertil Steril* 2011;**95**:2527–2530.
- Duffy JMN, Ahmad G, Mohiyiddeen L, Nardo LG, Watson A. Growth hormone for in vitro fertilisation. *Cochrane Database Syst Rev* 2010;**1**:CD000099.
- Duffy JMN, Arambage K, Correa FJS, Olive D, Farquhar C, Garry R, Barlow DH, Jacobson TZ. Laparoscopic surgery for endometriosis. *Cochrane Database Syst Rev* 2014;**4**:CD011031.
- Duffy JMN, Bhattacharya S, Herman M, Mol B, Vail A, Wilkinson J, Farquhar C. Reducing research waste in benign gynaecology and fertility research. *BJOG* 2017a;**124**:366–369.
- Duffy JMN, Hirsch M, Gale C, Pealing L, Kawsar A, Showell M, Williamson PR, Khan KS, Ziebland S, McManus RJ. A systematic review of primary outcomes and outcome measure reporting in randomized trials evaluating treatments for pre-eclampsia. *Int J Gynaecol Obstet* 2017c;**139**:262–267.
- Duffy JMN, Hirsch M, Kawsar A, Gale C, Pealing L, Plana MN, Showell M, Williamson PR, Khan KS, Ziebland S *et al.* Outcome reporting across randomised controlled trials evaluating therapeutic interventions for pre-eclampsia. *BJOG* 2017b;**124**:1829–1839.
- Duffy JMN, Hirsch M, Pealing L, Showell M, Khan KS, Ziebland S, McManus RJ. Inadequate safety reporting in pre-eclampsia trials: a systematic evaluation. *BJOG* 2017d. doi:10.1111/1471-0528.14969; in press.
- Duffy JMN, Johnson N, Ahmad G, Watson A. Postoperative procedures for improving fertility following pelvic reproductive surgery. *Cochrane Database Syst Rev* 2009;**1**:CD001897.
- Duffy JMN, McManus RJ. Influence of methodology upon the identification of potential core outcomes: recommendations for core outcome set developers are needed. *BJOG* 2016b;**123**:1599.
- Duffy JMN, Rolph R, Gale C, Hirsch M, Khan KS, Ziebland S, McManus RJ. Core outcome sets in women's and newborn health: a systematic review. *BJOG* 2017e;**124**:1481–1489.
- Duffy JMN, van 't Hooft J, Gale C, Brown M, Brogman W, Fitzpatrick R, Karumanchi SA, Lucas N, Magee L, Mol B *et al.* A protocol for developing, disseminating, and implementing a core outcome set for pre-eclampsia. *Preg Hyper* 2016a;**6**:274–278.
- Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Ytter Y, Glasziou P, DeBeer H *et al.* GRADE guidelines: 1. Introduction—GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;**64**:383–394.
- Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. *Br Med J* 2012;**344**:d7202.
- Hirsch M, Duffy JMN, Kuszniir JO, Davis CJ, Plana MN, Khan KS, International Collaboration to Harmonize Outcomes and Measures for Endometriosis. Variation in outcome reporting in endometriosis trials: a systematic review. *Am J Obstet Gynecol* 2016;**214**:452–464.
- Khalil A, Perry H, Duffy JMN, Reed K, Baschat A, Deprest J, Hecher K, Lewi L, Lopriore E, Oepkes D. Twin-Twin Transfusion Syndrome: a study protocol for developing, disseminating, and implementing a core outcome set. *Trials* 2017;**18**:325.
- Khan KS. The CROWN Initiative: journal editors invite researchers to develop core outcomes in women's health. *BJOG* 2014;**121**: 1181–1182.
- Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, Williamson PR. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *Br Med J* 2010;**340**: c365.
- Legro RS, Wu X, Barnhart KT, Farquhar C, Fauser BCJM, Mol B, Legro RS, Wu X, Barnhart K, Niederberger C *et al.* Improving the Reporting of Clinical Trials of Infertility Treatments (IMPRINT): modifying the CONSORT statement. *Hum Reprod* 2014;**29**:2075–2082.
- Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, Forbes A. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. *Cochrane Database Syst Rev* 2014;**1**:MR000035.
- Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, Williamson PR, Terwee CB. How to select outcome measurement instruments for outcomes included in a “Core Outcome Set” – a practical guideline. *Trials* 2016;**17**:449.
- Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. *BMJ* 2014;**349**:g6501.
- Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. *PLoS Med* 2011;**8**:e1000393.
- Smyth RMD, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: interviews with trialists. *Br Med J* 2011;**342**:c7153.
- Thoma ME, McLain AC, Louis JF, King RB, Trumble AC, Sundaram R, Buck Louis GM. Prevalence of infertility in the United States as estimated by the current duration approach and a traditional constructed approach. *Fertil Steril* 2013;**99**:1324–1331.e1.
- van 't Hooft J, Duffy JMN, Daly M, Williamson PR, Meher S, Thom E, Saade GR, Alfirevic Z, Mol BW, Khan KS. A core outcome set for evaluation of interventions to prevent preterm birth. *Obstet Gynecol* 2016;**127**:49–58.
- Whitehouse KC, Kim CR, Ganatra B, Duffy JMN, Blum J, Brahma D, Creinin MD, PePineres T, Gemzell-Danielsson K, Grossman D *et al.* Standardizing abortion research outcomes (STAR): a protocol for developing, disseminating and implementing a core outcome set for medical and surgical abortion. *Contraception* 2017;**95**:437–441.

- Wilkinson J, Roberts SA, Showell M, Brison DR, Vail A. No common denominator: a review of outcome measures in IVF RCTs. *Hum Reprod* 2016;**31**:2714–2722.
- Wilkinson J, Vail A, Roberts SA. Direct to consumer advertising of success rates for medically assisted reproduction: a review of national clinic websites. *BMJ Open* 2017;**7**:e012218.
- Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, Tugwell P. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;**13**:132.
- Webbe J, Brunton G, Ali S, Duffy JMN, Modi N, Gale C. Developing, implementing and disseminating a core outcome set for neonatal medicine. *BMJ Peds Open* 2017;**1**:e00048.
- Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, Sullivan E, Vanderpoel S, for the International Committee for Monitoring Assisted Reproductive and World Health Organization. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. *Fertil Steril* 2009;**92**:1520–1524.