Important research outcomes for treatment studies of perinatal depression: systematic overview and development of a core outcome set

C Hellberg, M Österberg, AK Jonsson, S Fundell, F Trönnerberg, M Jonsson, A Skalkidou

Published online: 3 May 2021. DOI: 10.1111/1471-0528.16780

Objective To develop a Core Outcome Set (COS) for treatment of perinatal depression.

Design Systematic overview of outcomes reported in the literature and consensus development study.

Setting International.

Population Two hundred and twenty-two participants, mainly patients, healthcare professionals and researchers, representing 13 countries.

Methods A systematic overview of outcomes reported in recently published research, a two-round Delphi survey and a consensus meeting at which the final COS was decided using modified nominal group technique.

Main results In the literature search, 1772 abstracts were identified and evaluated, and 165 studies were finally included in the review. In all, 106 outcomes were identified and included in the Delphi survey. In all, 222 participants registered for the first round of the Delphi survey and 151 (68%) responded. In the second round, 123 (55%) participants responded. Thirteen participants attended the consensus meeting, where the following nine outcomes were agreed upon for inclusion in the final COS: self-assessed symptoms of depression, diagnosis of depression by a clinician, parent to infant bonding, self-assessed symptoms of anxiety, quality of life, satisfaction with intervention, suicidal thoughts, attempted or committed suicide, thoughts of harming the baby, and adverse events.

Conclusions The relevant stakeholders prioritised outcomes and reached consensus on a COS comprising nine outcomes. We expect that this COS will contribute to the consistency and uniformity of outcome selection and reporting in future clinical trials involving treatment of perinatal depression.

Keywords Antenatal depression, perinatal depression, postpartum depression, patient and public involvement.

Introduction

Perinatal (or peripartum) depression refers to depression experienced during pregnancy (antenatal or prenatal depression) or after childbirth (postnatal or postpartum depression). The condition affects more than 10% of mothers and refers to depressive episodes starting in pregnancy or postpartum. Not only is the mother affected, but also the mother–infant attachment and bonding, as well as the cognitive skills of the children. In addition to these unwanted sequelae, the co-parent is also impacted and at increased risk of developing depression. Perinatal depression also comes at high societal cost. Perinatal depression is a complex heterogeneous disease with social and biological correlates, but its pathophysiology remains unclear. Recently, differential symptom phenotypes, possible...
trajectories\textsuperscript{11} and distinct pathophysiological pathways\textsuperscript{12} have been identified. Treatment options include pharma-cotherapy, psychotherapy and, in severe cases, electrocon- vulsive therapy.\textsuperscript{13}

However, in clinical trials involving this population there are inconsistencies in outcome selection and reporting. This raises concerns about possible outcome selection bias, hinders research synthesis and limits the potential to combine the findings of individual studies into summary estimates. One way to overcome this is to develop a core outcome set (COS); to date, none is available for perinatal depression.\textsuperscript{14} The issue of little information on a number of important outcomes has been raised by systematic review authors in relation to perinatal depression.\textsuperscript{15}

A COS is the minimum set of outcomes that should consistently be measured and reported in all clinical trials (and other studies). This does not restrict researchers from adding additional outcomes of relevance to their particular study. The Core Outcome Measures in Effectiveness Trials (COMET) initiative aims to standardise outcome reporting in trials, facilitates participation of diverse experts undertaking research and minimises duplication of work. A minimum set of outcomes is expected to provide greater uniformity of reporting in clinical trials and more data to inform future meta-
analyses.

The aim of this study was to develop a COS for clinical trials evaluating the effect of treatments for perinatal depression.

Methods

The project was registered in the COMET initiative registry (COMET registration number 1421) and has been developed in accordance with the COMET handbook,\textsuperscript{16} the COS STAndards for Development (COS-STAR) recommendations\textsuperscript{18} and the COS STAndards for Reporting (COS-STAR) recommendations.\textsuperscript{18} The project followed an a priori established protocol (Appendix S1) available on the website of the Swedish Agency for Health Technology Assessment and Assessment of Social Services (SBU) during the study period. The National Research Ethics Committee was consulted and concluded that the study did not require ethical approval.\textsuperscript{19}

Systematic overview

The following databases were searched: Medline (OVID), Cochrane Library (Wiley), PsycINFO (EBSCO), Cinahl (EBSCO) and ClinicalTrials.gov.

The full search strategy is presented in Appendix S2.

The following criteria were used to determine inclusion of the studies:

- Population: Either pregnant women or their partners suffering from depression (antenatal depression) or newly delivered mothers or their partners suffering from depression (postpartum depression).
- Intervention: Any intervention for depression
- Control: Any
- Outcome: Outcomes relating to the effect of treatment.
- Study design: Randomised controlled trial (RCT) or systematic reviews. Both protocols and published studies
- Publication date: January 2018 until October 2019 with the exception of ClinicalTrials.gov where no time restrictions were made

Exclusion criteria for studies were: Conference abstracts, studies investigating prevention of depression, studies investigating treatment of people suffering from a combination of depression and other chronic conditions such as HIV infection, and articles published in languages other than English or the Scandinavian languages.

Two authors (CH and MO) independently assessed the abstracts using the Rayyan platform.\textsuperscript{20}

The full-text papers were assessed independently by two authors (CH and MO) and any disagreement on the eligibility of included studies was resolved through discussion.

The following data were extracted from the studies:

- Reference information
- Study design
- Publication year
- Intervention
- Outcomes
- Measurement instruments
- Time for outcome measure

Data were extracted by one author and checked by another author. Any disagreement was resolved through discussion. After extraction, all outcomes that were unique were listed. Some very similar outcomes in the list were combined, for example outcomes measuring different hormonal or pharmaceutical levels were combined as biological parameters. The participants in the Delphi survey were also given the opportunity to add outcomes in the first survey round; however, we did not undertake any qualitative research with patients or any other of the stakeholder groups.

All unique suggested outcomes were added if they were considered to be outcomes and not background information (e.g. demographics and previous history of depression were regarded as background information).

Participants in the Delphi study

Invitation emails were sent to different relevant stakeholders such as Swedish and international user organisations, professional associations, researchers, Health Technology Assessment (HTA) organisations and other COS developers, as listed in Appendix S3. Information was also
disseminated via social media (Twitter, Facebook and LinkedIn) using SBU’s existing accounts as well as personal accounts of members of the project management team. Patients were mainly recruited through posts about the project on social media.

Those who wished to participate registered on SBU.se with the following information: name, email address, phone number (optional), stakeholder group, occupation (if relevant) and country of residence. After registration, participants were sent additional information about the study and about COS by email (Appendix S4).

Delphi method
We conducted an online two-round Delphi survey using the Delphi software.21 The survey was available in English or Swedish. The survey was piloted and adjusted by the study management group, involving professional expertise as well as patient perspectives, before being used.

Consensus definitions
Consensus definitions were set a priori as described:

- Consensus for an outcome to be included in COS (‘consensus in’): 70% or more of each stakeholder group scoring 7–9 AND less than 15% of participants scoring 1–3. If more than ten outcomes are scored as ‘consensus in’, prioritisation of which to include in the final COS will be done during the consensus meeting.
- Consensus for an outcome to be excluded from COS (‘consensus out’): 70% or more of each stakeholder group scoring 1–3 AND less than 15% of participants scoring 7–9.

No consensus: The outcomes were brought forward to the next survey.

In order to facilitate implementation and use of the COS, a pre-set goal of the consensus meeting was that the COS would comprise no more than ten outcomes. There is no consensus regarding a maximum number of outcomes to include in a COS, but a COS with fewer outcomes may be more feasible to implement. The decision to put a limitation on the number of outcomes resulted from previous SBU experience in other COS development, as well as experience from using existing COS when conducting systematic reviews.

A deviation from the protocol was made before the consensus meeting, because no outcomes were designated ‘consensus out’ during the Delphi process. To facilitate the consensus meeting and fruitful discussions, the project management team took the decision to continue with outcomes using at least one of the following three criteria:

- The outcomes scored as ‘consensus in’ after the Delphi survey
- The top ten outcomes from all four different stakeholder groups described below
- Outcomes considered critically important by 70% or more of one or more of the different stakeholder groups

Round 1 and 2
Participants were encouraged to complete the Delphi survey in each round. A maximum of three email reminders was sent to anyone failing to respond before the end of each round.

In the first round of the survey, all the outcomes identified were presented to the participants. Using a nine-point Likert scale, they were asked to rate the importance of inclusion of each outcome in the COS: (1–3: not important for inclusion; 4–6: important, but not critical for inclusion; 7–9: the outcome is of critical importance for inclusion). Participants were also invited to suggest additional relevant outcomes using free-text responses.

All stakeholders were grouped into four broader groups: people with personal experience of the condition/ or their relatives, clinicians, researchers and others (including policy-makers and HTA bodies). Descriptive statistics were used to summarise the results from Round 1. The results of each stakeholder group and the results of the total group were sent to each study participant. As no outcome was scored as ‘consensus out’, all outcomes from Round 1 and the additional suggested outcomes were included in Round 2. All respondents to Round 1 were invited to participate in Round 2 and asked to re-rate the outcomes. Using the criteria described above, 23 outcomes were presented by email to the representatives at the consensus meeting.

Consensus meeting
The consensus meeting involved 13 participants, including representatives from each stakeholder group (Table 2). The meeting opened with an initial briefing on the purpose and scope of the meeting. The meeting was chaired by an experienced facilitator from SBU and conducted in two sessions. SBU’s role during the workshops was to organise and facilitate discussions and make sure everyone was heard, but not to actively participate in the discussions.

The initial session consisted of group discussions in three small subgroups, comprising both patients and professionals, to achieve balance. Discussions were then held using a modified nominal group technique22 (with the ambition to take everyone’s opinions into account, as opposed to traditional voting, where only the largest group is considered) in the smaller groups first: discussions were thereafter held in the whole group. The participants discussed each outcome brought forward from Round 2. During discussion in the smaller groups, outcomes were sorted into three categories: (1) outcomes that should be included in the final COS; (2) outcomes where opinion was divided; and (3) outcomes which should not be included in the final COS. If necessary, the group was allowed to group or rename
outcomes if they believed that doing so could facilitate dissemination and usefulness. To begin the process, each individual presented their choices, including a short justification, to their subgroup. Subsequently, the members of the subgroup worked together to build a consensus that would best represent the views of the group. At the end of the exercise, the facilitators summed up decisions made by the three groups.

The second session consisted of a plenary discussion involving the entire group: the goal was to arrive at consensus as to which outcomes to include in the final COS.

Comments from other participants
The final COS was sent to all participants answering both Delphi rounds, in order to give all participants, the opportunity to comment on the results.

Results
Systematic overview
A total of 1772 abstracts were identified and evaluated, 284 papers/protocols were assessed in full and 165 studies were finally selected for inclusion in this review (Flowchart available in Figure S1). The included studies are presented in more detail in Table S1 and the excluded studies, with reasons for exclusion, in Table S2. Most of the included studies were protocols for RCTs (Table 1). On average, the RCTs contained six outcomes and the systematic reviews contained three outcomes (Table 1). The three most common outcomes in the included studies were self-assessed symptoms of depression, clinical diagnosis of depression and self-assessed symptoms of anxiety. The studies had a range of outcomes from 1 to 24 (median 5) (Table 1).

Most of the interventions referred to some form of ‘psychotherapy’. Other common intervention categories were ‘drugs’ and ‘complementary medicine’ (Table S1). Most of the studies, but not all, included at least one measurement related to depression. The range of different outcomes included in the same study to measure different aspects of depression ranged from 0 to 8.

Delphi study
A flow chart describing the steps in the development of the COS is presented in Figure 1. Following data extraction, 945 outcomes were identified. This number nevertheless includes variables that referred to the same outcome: for example, the level of depression assessed by the patient was identified 188 times in 133 studies. After extracting unique outcomes and combining similar outcomes, 93 outcomes remained (Figure 1). Five additional outcomes were suggested by the study management group and are marked in Table S3 where all outcomes are presented. This resulted in 98 outcomes in survey one.

A total of 222 individuals registered to participate in the survey, representing 13 countries and four continents (Table 2). The majority were health professionals (45%) and the smallest stakeholder group were HTA agencies, policy-makers and others (6%). Only one person representing a low-income country participated, while the other participants were from high-income countries (Table 2). No additional demographic data were collected. No fixed panel size was decided on a priori. As we analysed the different stakeholder groups separately, we do not believe that the difference in total numbers between the different groups has influenced the result. The panel members of the consensus meeting were decided a priori to be approximately 12–20, based on experience from previous meetings with similar design and aim.

One hundred and fifty-one (68%) responded in the first round, (Table 2). The distribution of answers throughout the stakeholder groups in each of the two rounds is presented in Table 2. A further seven outcomes were suggested by study participants and were included in Round 2 (Table S3).

Round 2 included 105 outcomes and was completed by 123 participants (55%). After the first two rounds, three outcomes scored as ‘consensus in’ were brought forward to the consensus meeting (Table S3).

The result for all outcomes in Rounds 1 and 2 for the whole group is presented in Table S1. The results according to stakeholder groups are also presented in Table S4 (survey 1) and Table S5 (survey 2).

Consensus meeting
There were 13 participants in the consensus meeting (15 had been invited, but two were unable to attend) (Table 2).

The final COS decided on after the consensus meeting comprised the following nine outcomes.

- Self-assessed symptoms of depression. Should be assessed with a scale that captures differences in sleep patterns
- Diagnosis of depression by a clinician, should include a structured interview
- Parent to infant bonding
- Self-assessed symptoms of anxiety
- Quality of life
- Satisfaction with intervention
- Suicidal thoughts, attempted or committed suicide
- Thoughts of harming the baby, including thoughts of extended suicide
- Adverse event

The final COS was sent to all participants answering both Delphi rounds, with the option for additional comments, no comments were received.

One of the decided outcomes, adverse events, is to be considered a comprehensive outcome. The main reason for
keeping this outcome broad and not specifying it further is the dependence on the type of intervention to be studied and the study period. This outcome will be different if the study only focuses on antenatal depression or on postnatal depression and needs to be specified in advance for each individual trial.

**Discussion**

**Main findings**

This study used robust methods to develop the first COS relevant to treatment of perinatal depression. In total, 93 unique outcomes were identified in the initial systematic
After a two-round Delphi survey, followed by a consensus meeting using a modified nominal group technique, our final COS comprised nine outcomes. With reference to the frequency of the selected nine outcomes in current research, only one, self-assessed symptoms of depression, appears in more than 50% of the identified studies. The second most frequently used outcome, diagnosis of depression by a clinician, occurred in only 33% of the studies (Figure 2). One of the outcomes, thoughts of harming the baby, including thoughts of extended suicide, did not occur in any of the identified studies. This shows the importance of enabling participants to suggest new outcomes that they consider important.

**Strengths and limitations**

The strengths of the study include the use of robust methods in COS development, including adherence to the COS-STAR statement, the thoroughness of the systematic review, the high number of participants and the diversity of stakeholders participating at each stage of the process. The study included patient representation, not only as participants in all steps of the process, but also in the project management group.

We sought international participation to ensure that the COS would have global relevance but were unfortunately not able to include an international panel in the consensus meeting. To anchor the COS suggested by the consensus meeting, all participants were given the opportunity to comment on the results. Despite aiming for an international audience, the majority of the respondents came from Sweden. We chose not to collect any additional demographic background information from the participants because that type of data would be classified as sensitive personal data, which could, in its turn, have affected the choice to participate for some of the patient participants.

Of importance, the consensus meeting included representatives from a variety of health professions/specialties and included women with experience of prenatal depression. This stakeholder composition permitted the health professionals, researchers and patients to bring their experience and perspectives to the issues under discussion. Each participant was able to gain a better understanding not only of what was important to other groups, but also what it was feasible to measure in all studies. Ultimately, this resulted in shared decision-making in a study that will impact future research.

There are some limitations. One is the time frame for the search in the systematic overview, which is limited to 20 months when it comes to published studies. However, we additionally reviewed all trials registered in ClinicalTrials.gov regardless of date and allowed for participants to...
suggest additional outcomes not identified in the literature. A scoping approach when it comes to finding outcomes has also been suggested in the literature.\(^{23}\) We believe that this was an effective strategy to reach outcome saturation without being too resource intensive. The contribution of unique outcomes derived mainly from study records in the ClinicalTrials.gov database, as well as outcomes added by participants in the first Delphi round.

In regard to the COS development one limitation is the pre-specified 70%/15% consensus definition used in this study, also commonly used by other COS developers.\(^{14,24}\) In our experience, people are very hesitant to score any outcome as of low importance and this limited the number of outcomes that were considered to be less important. It may be more beneficial in future studies to redefine the criteria for ‘consensus out’ during the Delphi surveys. In order to carry out our study, an adjustment was made, not pre-specified in the protocol, as to which outcomes to bring forward to the consensus meeting. This adjustment was made after careful consideration and discussion among all members of the project management group. The outcomes discussed in the consensus meeting were those with the highest ranking from each of the stakeholder groups, thus representing the commonly shared opinion of the participants in the survey.

Another potential limitation of this study is the large number of items to be scored in the Delphi surveys, which may have impacted negatively on response rates, which were 68% in the first survey and 55% in the second survey. Although this can be considered low, it did not differ among the different perspectives. Looking at other COS developed within the area of pregnancy and maternity care, response rates from all surveys vary substantially from as low as 20% up to 76%.\(^{25,26}\)

In the present paper, we present an overview of outcomes used in recent RCTs and systematic reviews as well as the development of a COS on what to measure in this research area. However, this study does not include guidance on how and when these outcomes should be measured, and we believe that a future project on that topic would increase the usability and implementation opportunities of the current COS.

The number of outcomes included in the final COS could be regarded as both a strength and a limitation. In order to facilitate the implementation of the COS, it was pre-specified as less than ten; the number finally included was nine. However, even a COS comprising nine outcomes may prove cumbersome if researchers want to be able to add additional research specific outcomes.

**Interpretation**

This is the first study to outline a COS for treatment of perinatal depression. The application of agreed methods in developing this COS and the participation of multiple stakeholder groups from several countries assure international applicability.

We encourage all investigators undertaking research in this field to report, as a minimum, this COS, in order to facilitate comparison among studies and to increase the potential for evidence synthesis across clinical studies. This will ultimately lead to improvement in the quality of research and delivery of evidence-based health care within this field.

However, while mandatory collection and reporting of all outcomes in the COS is recommended, researchers are still free to record any additional outcomes required for their study.

**Conclusion**

Relevant stakeholders agreed on the following outcomes for inclusion in the COS: self-assessed symptoms of depression, diagnosis of depression by a clinician, parent to infant bonding, self-assessed symptoms of anxiety, quality of life,
satisfaction with intervention, suicidal thoughts, attempted or committed suicide, thoughts of harming the baby, and adverse events. We expect this COS will help bring consistency and uniformity to outcome selection and reporting in future clinical trials involving treatment of perinatal depression.

Disclosure of interests
The authors report no conflict of interests; all authors as well as those included in the COS development consensus meeting filed a conflict of interests form used by Swedish governmental agencies before engagement. These are available upon request. Completed disclosure of interests form available to view online as supporting information.

Contribution of authorship
CH, MÖ, AS, MJ and FT contributed to study concept and design, analysis and interpretation of data and preparation of materials for COS participants. AJ performed the literature search. CH and MÖ selected the studies and extracted the relevant information. MJ and AS contributed as academic and clinical experts and FT was responsible for the patient perspective. Design and conduction of the Delphi surveys were by SF and design and conduction of consensus meeting were by CH, MÖ, AS, MJ and FT. All material that was sent out to the participants was reviewed by FT in terms of how easy they were to understand and read. CH and MÖ drafted the manuscript. Critical revision of the manuscript for important intellectual content was by the entire study management group.

Details of ethics approval
The National Research Ethics Committee was consulted and concluded that the study did not require ethical approval.16

Funding
The project was conducted within the Swedish Agency for Health Technology Assessment and Assessment of Social Services assignment, external funding was not sought or used.

Acknowledgements
We thank all the stakeholders who participated in the Delphi study and consensus meeting.

Data availability statement
Data are available on request.

Supporting Information
Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1. Flow chart of the systematic overview of outcomes.

Table S1. Included studies.
Table S2. Excluded studies.
Table S3. Summary of results from Surveys 1 and 2. Outcomes discussed during the consensus meeting are highlighted in bold.
Table S4. Results from Survey 1.
Table S5. Results from Survey 2.
Appendix S1. Protocol.
Appendix S2. Search strategies.
Appendix S3. Organisations contacted with information about the project and invitation to participate
Appendix S4. Information to the participants.

References
1 American Psychiatric Association. DSM-5 Task Force. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders: DSM-5. 5th edn. Arlington, VA: American Psychiatric Association; 2013.
2 Woody CA, Ferrari AJ, Siskind DJ, Whiteford HA, Harris MG. A systematic review and meta-regression of the prevalence and incidence of perinatal depression. J Affect Disord 2017;219:86–92.
3 Agnafors S, Sydsjö G, Dekyser L, Svedin CG. Symptoms of depression postpartum and 12 years later-associations to child mental health at 12 years of age. Matern Child Health J 2013;17:405–14.
4 Duan C, Hare MM, Staring M, Deligiannidis KM. Examining the relationship between perinatal depression and neurodevelopment in infants and children through structural and functional neuroimaging research. Int Rev Psychiatry 2019;31:264–79.
5 Oh Y, Joung Y-S, Baek JH, Yoo N. Maternal depression trajectories and child executive function over 9 years. J Affect Disord 2020;276:646–52.
6 Smith A, Twynstra J, Seabrook JA. Antenatal depression and offsping health outcomes. Obstet Med 2019;13:55–61.
7 Paulson JF, Bazemore SD. Prenatal and postpartum depression in fathers and its association with maternal depression: a meta-analysis. JAMA 2010;303:1961–9.
8 Bauer A, Knapp M, Parsonage M. Lifetime costs of perinatal anxiety and depression. J Affect Disord 2016;192:83–90.
9 Santos H Jr, Tan X, Salomon R. Heterogeneity in perinatal depression: how far have we come? A systematic review. Arch Womens Ment Health 2017;20:11–23.
10 Putnam KT, Wilcox M, Robertson-Blackmore E, Sharkey K, Bergink V, Munk-Olsen T, et al. Clinical phenotypes of perinatal depression and time of symptom onset: analysis of data from an international consortium. Lancet Psychiatry 2017;4:477–85.
11 Wikman A, Axford C, Iliadis SI, Cox J, Fransson E, Skalkidou A. Characteristics of women with different perinatal depression trajectories. J Neurosci Res 2020;98:1268–82.
12 Meltzer-Brody S. New insights into perinatal depression: pathogenesis and treatment during pregnancy and postpartum. Dialogues Clin Neurosci 2011;13:89–100.
13 Pearlstein T. Perinatal depression: treatment options and dilemmas. J Psychiatry Neurosci 2008;33:302–18.
14 SBU. Core Outcome Sets for Research within the Area of Maternity Care. Overview of Completed and Ongoing Studies. Stockholm, Sweden: SBU – Swedish Agency for Health Technology Assessment and Assessment of Social Services; 2020. Report no 309 [https://www.sbu.se/309e]. Accessed 8 September 2020.
Treatment of perinatal depression: a core outcome set

15 Molyneaux E, Howard LM, McGeown HR, Karia AM, Trevillion K. Antidepressant treatment for postnatal depression. Cochrane Database Syst Rev 2014;9:CD002018.

16 Williamson PR, Altman DG, Bagley H, Barnes KL, Blazey JM, Brookes ST, et al. The COMET handbook: version 1.0. Trials 2017;18:280.

17 Kirkham JJ, Davis K, Altman DG, Blazey JM, Clarke M, Tunis S, et al. Core outcome Set-STAndards for development: the COS-STAD recommendations. PLoS Med 2017;14:e1002447.

18 Kirkham JJ, Gorst S, Altman DG, Blazey JM, Clarke M, Devane D, et al. Core outcome Set-STAndards for reporting: the COS-STAR statement. PLoS Med 2016;13:e1002148.

19 Ethical advisory opinion: Swedish Ethical Review Authority, reference dnr 2019-05673 (December 18, 2019).

20 Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016;5:210.

21 Defgo.com [Internet]. Copenhagen: defgo; c2020 [cited 2020 April 1]. Available from: https://www.defgo.com/se/.

22 Harvey N, Holmes CA. Nominal group technique: an effective method for obtaining group consensus. Int J Nurs Pract 2012;18:188-94.

23 Beune IM, Ganzvoort W, Gordijn SJ. Core outcome sets are valuable, but methodological evidence can improve robustness. BJOG 2020;127:1527.

24 Williamson PR, Altman DG, Bagley H, Barnes KL, Blazey JM, Brookes ST, et al. The COMET handbook: version 1.0. Trials 2017;18:280.

25 Bogdanet D, Reddin C, Macken E, Griffin TF, Fheleboom N, Biesty L, et al. Follow-up at 1 year and beyond of women with gestational diabetes treated with insulin and/or oral glucose-lowering agents: a core outcome set using a Delphi survey. Diabetologia 2019;62:2007-16.

26 van ‘t Hooft J, Duffy JMN, Daly M, Williamson PR, Meher S, Thom E, et al. A core outcome set for evaluation of interventions to prevent preterm birth. Obstet Gynecol 2016;127:49-58.