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

Christel Hellberg[1], Marie Österberg[1], Ann Kristine Jonsson[1], Sara Fundell[1], Frida Trönnberg[2], Maria Jonsson[3], Alkistis Skalkidou[3]

[1] Swedish Agency for Health Technology Assessment and Assessment of Social Services, Stockholm, Sweden

[2]Patient representative

[3] Dept. of Women's and Children's Health, Uppsala university, Sweden

Short title: Treatment of perinatal depression: a core outcome set

Abstract

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 using a Delphi survey and modified nominal group technique.

Setting

International.

Population

Two hundred and twenty-two participants, representing thirteen countries.

Methods

A systematic overview of outcomes reported in recently published research, a two-round Delphi survey, a consensus meeting at which the final COS was decided.

Main results

In the literature search, 1772 abstracts were identified and evaluated, 284 papers/protocols were assessed in full and 165 studies were finally included in the review. In all, 106 outcomes were identified and thus included in the Delphi survey. 222 participants registered for the first round of the Delphi survey and 151 (68%) responded. In the second round, 123 (55%) participants responded. The following 9 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 hope that this COS will contribute to consistency and uniformity of outcome selection and reporting in future clinical trials involving treatment of perinatal depression

Funding

This article is adapted from a report by SBU, which provided funding for the study.

Keywords: perinatal depression, postpartum depression, antenatal depression

Tweetable abstract

Development of a core outcome set regarding treatment for perinatal depression by @SBU_en

Corresponding author

Christel Hellberg, Ph.D., Swedish Agency for Health Technology Assessment and Assessment of Social Services, Stockholm, Sweden. Christel.hellberg@sbu.se

Abbreviations

COMET	Core Outcome Measures in Effectiveness Trials Initiative
COS	Core Outcome Set/Sets
COS-STAD	The Core Outcome Set-STAndards for Development
RCT	Randomised Controlled Trial
SBU	Swedish Agency for Health Technology Assessment and Assessment of Social Services
HTA	Health Technology Assessment

Introduction

Perinatal (or peripartum) depression refers to depression experienced during pregnancy (antenatal or prenatal depression) or after childbirth (postnatal or postpartum depression) (1). The condition affects more than 10% of mothers (2) and refers to depressive episodes starting in pregnancy or postpartum (1). Not only is the mother affected, but also the mother-infant attachment and bonding as well as the cognitive skills of the children (3-6). In addition to these unwanted sequelae, the other parent is also impacted and at increased risk of developing depression (7). Perinatal depression also comes at high societal cost (8). Perinatal depression is a complex heterogeneous disease (9) with social and biological correlates, but its pathophysiology remains unclear. Recently, differential symptom phenotypes. (10), possible trajectories (11) and distinct pathophysiological pathways (12) have been identified. Treatment options include pharmacotherapy, psychotherapy, and in severe cases, electroconvulsive therapy (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 (14).

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 and has been developed according to the COS-STAD Recommendations (15). 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 (16).

The three parts of this study were: 1) a systematic overview of the outcomes reported in recently published systematic reviews and RCTs as well as in relevant ongoing RCTs 2) a two-round online Delphi survey in which all outcomes were scored and prioritised by key stakeholder groups, to provide a preliminary list of final outcomes; and 3) a consensus meeting where the final list of outcomes was decided.

Systematic overview

The following databases were searched for systematic reviews or randomized controlled trial: Medline (OVID), Cochrane library (Wiley), PsycINFO (EBSCO), Cinahl (EBSCO) and ClinicalTrials.gov.

The search was restricted to 2018 until October 2019, with the exception of ClinicalTrials.gov were no time restrictions were made. No limitation to language was made but only studies in English or the Scandinavian languages were included.

The full search strategy is presented in Appendix s2.

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

Inclusion criteria:

- Population: Pregnant women or their partners suffering from depression or new mothers or their partners suffering from depressions
- Intervention: Any intervention for depression
- Control: Any type of control
- Outcome: Outcomes relating to the effect of treatment
- Study design: Randomised controlled trials or systematic reviews. Both protocols and published studies were included

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

Two review authors (CH and MÖ) independently assessed the abstracts using the Rayyan abstract screening application (17). All potentially relevant studies and those appearing to meet the inclusion criteria, or for which there were insufficient data in the title and abstract to make a clear decision, were retrieved in full text.

The full-text papers were assessed independently by two review authors (CH and MÖ) 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/s used
- Outcomes, including whether these stated as primary or secondary
- Measurement instruments used
- Time for outcome measure

Data were extracted by one review author and checked by another review author. Any disagreement was resolved through discussion. At this stage, experts from the study management group also had the opportunity to add outcomes they considered important, but which did not appear in the literature (AS, MJ and FT). 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. 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, different professional associations, Health technology assessment (HTA) organisations, other COS developers in the field, etc. All participants were asked to forward the invitation to others whom they regarded as having the required expertise. Information was also disseminated via social media (Twitter, Facebook and LinkedIn).

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

Delphi method

We conducted an online two-round Delphi survey using the Defgo software (18). The survey was available in English or Swedish.

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 to 9 AND <15% participants scoring 1 to 3. If more than 10 outcomes are scored as “consensus in”, prioritisation of which to include in the final core outcome set 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 to 3 AND <15% of participants scoring 7 to 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 10 outcomes.

A deviation from the protocol was made before the consensus meeting, as 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 the following 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 questionnaire in each round (information given to the participants as well as a plain language explanation of all the outcomes included in the survey is available in Appendix S3). 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 in the systematic review, grouped by domain, 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: is 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 (no limit to the number of outcomes suggested) using free-text responses.

All stakeholders were grouped into four broader groups, i.e. 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. The meeting opened with an initial briefing on the purpose and scope of the meeting. The results of the systematic overview and Delphi survey were presented. The meeting was chaired by an experienced facilitator from SBU and conducted in two sessions. SBU’s role during the workshops was to organize and facilitate discussions, but SBU members did not actively participate in the discussions.

The initial session consisted of group discussions in three small subgroups, comprising both patients and professionals, to achieve balance. The facilitators moderated the discussions (one in each subgroup). Facilitators ensured that everyone had the opportunity to be heard, took notes from the discussions and documented the decisions. For the discussions, each outcome was written on separate A4-cards. The results from the Delphi Round 2 for that specific outcome by stakeholder group were printed on the back of the cards. Discussions were then held using a modified nominal group technique 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 which should be included in the final core outcome set; (2) outcomes where opinion was divided; and (3) outcomes which should not be included in the final core outcome set. 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. The participants gathered around an area where the A4-cards had been distributed by the facilitators according to a "diamond-shape," based on the grouping made in the first session, with outcomes which all three subgroups considered should be included in the COS at the top, outcomes where groups had different views in the middle and at the bottom outcomes not considered by any group as important enough to be included. The members then moved the cards around during the discussion and finally agreed on 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 (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 RCT:s contained six outcomes and the systematic reviews four 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 0-8.

Delphi study

A flow chart describing the different steps in the development of the COS is presented in Figure 1. Following data extraction, 945 outcomes were identified. This number nevertheless includes variables which 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; considered by the group to be important but lacking in studies published to date. This resulted in 98 outcomes in survey one. In the first round of the survey, seven additional outcomes were suggested, leaving 105 outcomes for Round 2.

Two hundred and twenty-two people registered to participate in the survey, representing 13 countries and 4 continents. 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 7 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 (n = 3):

- Self-assessed symptoms of depression (should be assessed with a scale that captures differences in sleep-patterns)
- Suicidal thoughts, attempted or committed suicide
- Thoughts of harming the baby, including thoughts of extended suicide

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 2 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. Includes spontaneous or induced abortion, miscarriage, fetal death and death during the first week of life

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

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 overview. 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 which 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-STAD 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. In order to anchor the COS suggested by the consensus meeting, all participants were given the opportunity to comment on the results.

Of importance, the COS consensus meeting included representatives from a variety of health professions/specialties and included women with experience of prenatal depression. This stakeholder composition permitted both 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 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 to our study. One is the prespecified 70/15% consensus definition used in this study, also commonly used by other COS developers (14, 19). 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 prespecified 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 rounds 1 and 2 of the online Delphi survey, which may have impacted negatively on the survey response rate.

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 prespecified as less than 10; the number finally included was nine. However, a COS comprising even 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 therefore 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 healthcare 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

Using consensus development methods, relevant stakeholders agreed on the following nine 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 hope that this COS will help bring consistency and uniformity to outcome selection and reporting in future clinical trials involving treatment of perinatal depression.

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

Contribution of authorship: Study concept and design: CH, MÖ, AS, MJ, FT. Literature search AJ. Selection of studies and extraction of the relevant information CH and MÖ. Analysis and interpretation of data and preparation of materials for COS participants: CH, MÖ, AS, MJ, FT. Design and conduction of the delphi surveys SF. Design and conduction of consensus meeting: CH, MÖ, AS, MJ, FT Drafting of the manuscript: CH and MÖ. Critical revision of the manuscript for important intellectual content: AS, MJ, FT, CH, MÖ, SF, AJ.

Data availability Data are available on request.

Disclosure of interests: The authors report no conflict of interest, all authors as well as those included in the COS development consensus meeting filed a conflicts of interest form used by Swedish governmental agencies before engagement. These are available upon request.

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.

References

1. American Psychiatric Association. DSM-5 Task Force.American Psychiatric Association. Diagnostic and statistical manual of mental disorders : DSM-5. 5. ed. 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, Dekeyser 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(3):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(3):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 offspring health outcomes. Obstet Med. 2019;13(2):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(19):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(1):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(6):477-85.

11. Wikman A, Axfors C, Iliadis SI, Cox J, Fransson E, Skalkidou A. Characteristics of women with different perinatal depression trajectories. J Neurosci Res. 2020;98(7):1268-82.
12. Meltzer-Brody S. New insights into perinatal depression: pathogenesis and treatment during pregnancy and postpartum. Dialogues Clin Neurosci. 2011;13(1):89-100.
13. Pearlstein T. Perinatal depression: treatment options and dilemmas. J Psychiatry Neurosci. 2008;33(4):302-18.
14. SBU. Core outcome sets for research within the area of maternity care. Overview of completed and ongoing studies. Stockholm: SBU – Swedish Agency for Health Technology Assessment and Assessment of Social Services; 2020. Report no 309. [cited 2020 Sep 8]. Available from: https://www.sbu.se/309e.
15. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. PLOS Medicine. 2017;14(11):e1002447.
16. Swedish Ethical Review Authority, reference dnr 2019-05673.
17. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Systematic Reviews. 2016;5(1):210.
18. Defgo [Internet]. Defgo. [cited 2020 April 1]. Available from: https://www.defgo.com/se/.
19. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials 2017;18:280.