



Published in final edited form as:

Obstet Gynecol. 2016 January ; 127(1): 49–58. doi:10.1097/AOG.0000000000001195.

A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth

Janneke van 't Hooft, James M. N. Duffy, Mandy Daly, Paula R. Williamson, Shireen Meher, Elizabeth Thom, George R. Saade, Zarko Alfirovic, Ben Willem J. Mol, Khalid S. Khan on behalf of On behalf of the Global Obstetrics Network (GONet)

Department of Obstetrics and Gynecology, Academic Medical Center, Amsterdam, the Netherlands; the Irish Neonatal Health Alliance, Wicklow, Ireland; Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, MRC North West Hub for Trials Methodology Research, Department of Biostatistics, University of Liverpool, and Liverpool Women's Hospital, Liverpool, and Imperial College Health Care NHS Trust, Queen Charlotte's and Chelsea Hospital, and the Women's Health Research Unit, the Blizard Institute, Barts and the London School of Medicine and Dentistry, London, United Kingdom; the George Washington University Biostatistics Center, Rockville, Maryland; the Department of Obstetrics and Gynecology, University of Texas Medical Branch Hospitals, Galveston, Texas; and the Department of Obstetrics and Gynaecology, the Robinson Research Institute, School of Paediatrics and Reproductive Health, University of Adelaide, Adelaide, Australia.

Abstract

Objective—To develop consensus on a set of key clinical outcomes for the evaluation of preventive interventions for preterm birth in asymptomatic pregnant women.

Methods—A two-stage web-based Delphi survey and a face-to-face meeting of key stakeholders were employed to develop consensus on a set of critical and important outcomes. We approached five stakeholder groups (parents, midwives, obstetricians, neonatologists and researchers) from middle and high-income countries. Outcomes subjected to the Delphi survey were identified by systematic literature review and stakeholder input. Survey participants scored each outcome on a 9-point Likert scale anchored between 1 (limited importance) and 9 (critical importance). They had the opportunity to reflect upon total and stakeholder sub-group feedback between survey stages. For consensus, defined *a priori*, outcomes required at least 70% of participants of each stakeholder group scoring them as 'critical' and less than 15% as 'limited'.

Results—A total of 228 participants from five stakeholder groups from three lower-middle-income countries, seven upper-middle-income countries and 17 high-income countries were asked to score 31 outcomes. Of these participants, 195 completed the first survey and 174 the second.

Corresponding author: Janneke van 't Hooft, Department of Obstetrics and Gynecology, Room H4-240, Academic Medical Center, PO Box 227700, Amsterdam 1105DE, the Netherlands; j.vanthooft@gmail.com.

Presented as a poster at the 4th meeting of the Core Outcome Measures in Effectiveness Trials (COMET) Initiative, November 19–20, 2014, Rome, Italy, and at the Global Obstetrics Networks annual meeting, February 7, 2015, San Diego, California.

Financial Disclosure

The authors did not report any potential conflicts of interest.

Trial registration: COMET Registration Number: 603 <http://www.comet-initiative.org/studies/details/603>

Consensus was reached on 13 core outcome: four related to pregnant women: maternal mortality, maternal infection or inflammation, prelabor rupture of membranes, and harm to mother from intervention. Nine related to offspring: gestation age at birth, offspring mortality, birth weight, early neurodevelopmental morbidity, late neurodevelopmental morbidity, gastrointestinal morbidity, infection, respiratory morbidity, and harm to offspring from intervention.

Conclusions—This core outcome set for studies that evaluate prevention of preterm birth developed with an international multidisciplinary perspective will ensure that data from trials that assess prevention of preterm birth can be compared and combined.

INTRODUCTION

Clinical trials, systematic reviews and guidelines evaluate interventions by comparing outcomes chosen to reflect beneficial and harmful effects. Systematic reviews have the potential to minimize bias and to increase the precision of measurements of treatment effects by quantitative pooling (meta-analysis) of similar clinical trial outcomes. However, this method does not work if clinical trials collect different outcomes. The lack of consistency in outcomes reported in comparative health research evaluating interventions for preterm birth has led to over 72 different primary outcomes being reported in 103 randomised trials.¹ Such heterogeneity results in substantial outcome reporting bias and an inability to synthesise results across studies in systematic reviews.² As a consequence, 33 Cochrane reviews on preterm birth have reported on 29 different primary outcomes.¹ This problem could be addressed by the use of a core outcome set, that is, a set of critical and important outcomes that should be measured and reported, as a minimum, in a standardised manner in research.³ Such a set currently does not exist and its need has recently being expressed in a systematic review of studies of preterm neonates. They found that the outcome “chronic lung disease,” considered an important outcome, was found to be missing in 55% of relevant systematic reviews.²

A core outcome set captures the key outcomes from those that could be or have been used in trials of a specific topic. These core outcomes sets should be included in future studies of that topic. This does not imply that a particular trial should be restricted to those outcomes in the core set. Ideally, core outcomes will always be collected and reported, but researchers will continue to explore other outcomes.³ In many trials, the primary outcome would be expected to be one of those contained in the core set, although this is not part of the definition of a core outcome set. Successful implementation of a core outcome set for rheumatoid arthritis has resulted in improved harmonization of research by establishing outcomes which are now more frequently measured by researchers.⁴

The aim of our project was to use robust consensus methods and engage all key stakeholders to identify a set of critical and important outcomes (core outcome set) for the evaluation of preventive interventions for the preterm birth in asymptomatic pregnant women.

MATERIALS AND METHODS

A protocol with explicitly defined objectives, formal consensus development methods, criteria for participant identification and selection, and statistical methods was developed.

The study was prospectively registered with the Core Outcome Measures in Effectiveness Trials (COMET) initiative (registration number 603 available online at www.comet-initiative.org/studies/details/603). The ethics board of the Academic Medical Center, Amsterdam, The Netherlands, advised that ethical approval was not required (reference number E2-172) because this project should be considered as service evaluation and development.

The target of the core outcome set was to capture important outcomes for individual studies, systematic reviews, and guidelines for preterm birth prevention in asymptomatic woman. For our purposes, preterm birth was defined as neonates born alive before 37 weeks of gestation.⁵ An asymptomatic woman was defined as one without symptoms of preterm labor (e.g. increased uterine contractions, menstrual cramps of backache, color change of vaginal discharge, prelabor rupture of membranes). Preventive treatment of preterm birth was defined as one started before any symptoms of preterm labor were present. This preventive strategy could be pharmacologic (e.g. progesterone, marine oils, probiotics) or non-pharmacologic (e.g. cerclage, pessary, lifestyle interventions and alternative therapies).

A Project Steering Committee was established to give guidance to the different phases of this project consisting of two obstetricians (Irene de Graaf, Khalid S. Khan), two neonatologists (Timo de Haan, Stephen Kempley), two midwives (Felipe Castro, Birgit van der Goes), two patient representatives (Aoife Ahern, Mandy Daly) and three methodologists with experience in formal consensus and/or core outcome set methods (James Duffy, Brent Opmeer, and Paula Williamson).

A systematic literature review was undertaken searching the Cochrane Pregnancy and Childbirth Group's (PCG) Trials Register.¹ The Pregnancy and Childbirth Group register is maintained by monthly searches of the Cochrane Central Register of Controlled Trials and weekly searches of EMBASE and MEDLINE and hand-searches of 30 journal and conference proceedings (from January 1997 to January 2011). The register was searched utilizing the register's codes for preterm birth. Two reviewers (S.M. and Z.A.) independently screened titles and abstracts. They critically reviewed the full text of selected studies and extracted reported outcomes. Any discrepancies were resolved by discussion. In addition, all delegates (n=168) of the First European Spontaneous Preterm Birth Congress (Svendborg, Denmark, May 24–25, 2014), mainly representing obstetricians and researchers, but also midwives, neonatologists and members of industry, were requested via e-mail to recommend potential outcomes.

Patient representatives and parents were invited through social media (Twitter and patient forums on Facebook) to participate in an online questionnaire to share their opinions regarding outcomes relevant to preterm birth. Members of patient organisations including the European foundation for the Care of Newborn Infants, their partner organizations, and parental forums of neonatal baby units were e-mailed by their own organization including an invitation for the online questionnaire through an electronic newsletter. Patients also contributed their opinions through in-person semistructured interviews conducted by one of the authors (J.v.t.H.).

The Project Steering Committee identified outcomes that were duplicated as a result of varied terminologies used by different stakeholders and for grouping closely related outcomes into overarching domains. This outcome inventory of 29 outcomes was entered into a Delphi process (Figure 1).

We used a two-round electronic Delphi survey design, a well-established method to elicit consensus based on an iterative process with anonymous consultation and with controlled feedback and quantified analysis of the responses.⁶ A priori we agreed the important methodological features for our Delphi process: [1] composition of the group; [2] anonymity; [3] how to assess the importance of outcomes; [4] method of feedback of results to participants; [5] how consensus would be reached; [6] how to assess possible attrition bias.

The setting for the Delphi survey was multinational involving stakeholders from middle- and high-income countries. A formal written invitation was e-mailed to all members of the Cochrane Pregnancy and Childbirth group (n=30), the Core Outcomes in Women's Health initiative (n=77), the European Preterm Birth Congress (n=168), and the Global Obstetrics Network (n=237). Most members of these organizations are researchers (methodologists), obstetricians (mainly specialized in maternal fetal medicine) or neonatologists. The European foundation for the Care of Newborn Infants approached their members themselves, including their partner organizations in Australia, Belgium, Bulgaria, Canada, Chile, Croatia, Cyprus, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Israel, Italy, Lithuania, Mexico, the Netherlands, Norway, Poland, Portugal, Spain, Turkey, United Kingdom, and the United States. All midwives from 'Barts Health Nursing and Midwifery' (n=132) and some midwives of the School of Nursing and Midwifery (Galway, Ireland) and the Dutch Consortium for Healthcare Evaluation in Obstetrics and Gynaecology were approached. With this approach we aimed to targeted midwives who were active in research (50%) and midwives who were not active in research (50%). In total 337 obstetricians, 152 midwives, 174 researchers, 75 neonatologists, and an unknown number of parents (through the previously mentioned patient organizations) were invited.

We used LimeSurvey for the Delphi survey. The survey was piloted first by eight people representing every stakeholder group. No changes were needed after the pilot. The official survey had a closing date of 5 weeks after the date of invitation for every Delphi round. An e-mail reminder was sent to participants on days 7, 14, 21, and 28. Nonresponders in the first round were not invited to participate in the subsequent round.

Participants were asked to rate the importance of each outcome on a 9-point Likert scale anchored between 1 ('limited importance') and 9 ('critical importance'). The scale is recommended by the Grading of Recommendations Assessment, Development and Evaluation working group: 1–3: limited importance; 4–6: important but not critical; 7–9: critical.⁷ Participants were invited to recommend additional potential outcomes for consideration at the end of the survey using free-text responses.

The individual, stakeholder group and total results from the first round were relayed back to participants by e-mail; the individual responses directly after filling in the first round

questionnaire, the stakeholder group, and total group responses were fed back anonymously 1 day prior to the invitation to the second round of the Delphi survey. Furthermore, participants of the second survey were able to see the mean value of the total group responses from the first Delphi round while completing the survey. Participants were asked to score all the individual outcomes again using the same 9-point Likert scale. No outcomes were excluded in this round to ensure a holistic approach to scoring in round 2.

The Delphi survey responses were analyzed using SPSS version 21.0. For each outcome the median and interquartile range were calculated. Frequency tables of all scores were generated, as well as boxplots for visualization (that were used to relay back the whole and stakeholder group responses). We defined consensus a priori. Core outcomes required at least 70% of participants in each stakeholder group scoring the outcome as 'critical' and less than 15% of participants in each stakeholder group scoring the outcome as 'limited importance'.⁸ Outcomes which should not be included in a core outcome set required at least 70% of participants in each stakeholder group scoring the outcome as 'limited importance' and less than 15% of participants in each stakeholder group scoring the outcome as 'critical'. If outcomes did not meet either criteria they were classified as outcomes with no consensus. Attrition bias (e.g. a selective group did not respond to the second round of the survey or a selective group participated in the consultation meeting) was assessed by 1) comparing the distribution of median first round scores across the outcomes for those not participating in the second round with those who did; and 2) comparing the distribution of median round 2 scores across the outcomes for those participating in the consultation meeting compared with those who did not.

The final phase of the study was a face-to-face consultation meeting with participants of the Delphi exercise representing all stakeholder groups (Washington, DC, November 9, 2014). This meeting was organized within a meeting for a prospective individual participant data analysis project for studies on the use of pessary in the prevention of preterm birth in asymptomatic women. Eleven participants of this prospective individual patient data project did also take part in the Delphi survey earlier. They mainly represented the stakeholder groups of obstetricians and methodologists. Representatives from the other stakeholder groups (parents, midwives and neonatologists), who were living close to the location of the consultation meeting, were invited for this consultation meeting as well. In total 23 obstetricians, 10 researchers, two neonatologists, two patient representatives, and one midwife were invited to attend this meeting. Information material on the purpose of the consultation meeting and the Delphi round 2 results were sent to participants before the meeting. A plenary presentation on the Delphi survey outcomes was complemented by small group sessions (mixed groups) where participants expressed their views on the candidate outcomes. Only outcomes that did not reach full consensus in the Delphi exercise were presented to the attendees of the meeting with an anonymous voting using electronic touchpads. Consensus in the consultation meeting required a majority of 70% of participants from each stakeholder group approving an individual outcome as 'critical' according to the 1–9 Likert scale. With the permission of the participants the consultation meeting was recorded.

RESULTS

The systematic review yielded 170 randomised trials and 33 reviews and protocols. The flowchart and more information on the selection process and the systematic review are reported elsewhere.¹ We identified 72 outcomes reported as primary outcomes and 155 outcomes reported as secondary outcomes. A further 25 outcomes were recommended by participants of the First European Spontaneous Preterm Birth Congress, and eight additional outcomes were recommended by patients through interview or online questionnaires (Figure 1).

The Project Steering Committee considered all 260 identified outcomes. The committee excluded 36 outcomes that were not relevant to the study's population, 92 outcomes that were rather outcome measurement instruments or definitions of a particular outcome, and 17 outcomes that were duplicates (Appendix 1). Subsequently, 86 different outcomes (with some closely related) were grouped into 29 outcome domains that were entered into the Delphi process (Figure 1).

In round 1 of the survey, overall, 195 (86%) of the 228 participants from five stakeholder groups representing three lower-middle-income countries, seven upper-middle-income countries and 17 high-income countries (classification according to http://data.worldbank.org/about/country-and-lending-groups#Lower_middle_income) responded (Table 1). The Project Steering Committee considered the free text responses of participants and entered an additional 2 outcomes (offspring circulatory morbidity and offspring metabolic morbidity) into the Delphi survey round 2 and considered changes in the formulation of some outcomes (Appendix 2)

Round 2 of the survey was completed by 174 participants (89% response rate). Participants reflected on the stakeholder group response and total group responses of the 31 outcomes included in round 2 (Table 2). They reached full consensus in all stakeholder groups on 11 outcomes (Appendix 3). They failed to reach consensus regarding the remaining 20 outcomes. Ten of the 20 outcomes that did not reach consensus in the Delphi survey were considered in the consultation meeting. These 10 outcomes were outcomes that came out of the Delphi survey as consensus 'in' (ie, greater than 70% of the stakeholder group scoring the outcome as 'critical') by at least one stakeholder group (n=9 outcomes) or were listed in the top 10 of most important outcomes (n=1).

Participants who did not respond to the second round Delphi survey scored comparable median scores in the first round survey (with overlap in interquartile ranges) when compared to the scores of those who participated in both surveys (Appendix 4). Also, the median second round scores of participants who attended the consultation meeting did not differ significantly from the median scores of those who did not attend this meeting.

At the stakeholder meeting in Washington, DC, 29 participants representing all stakeholder groups discussed and voted on the 10 outcomes that did not reach full consensus by all stakeholder groups in the Delphi exercise (i.e. the nine outcomes that were consensus in by 1 or more in the stakeholder group and the outcome that was listed in the top 10; Figure 1). Only the outcome birth weight was rated by greater than 70% of all stakeholder groups with

a Likert score of 7–9. Minutes of the discussion and arguments for including or excluding outcomes are provided in Appendix 5.

The Project Steering Committee considered the results of the Delphi method and consultation meeting. The committee discussed and ratified all the 12 selected core outcomes of the Delphi and consultation meeting process. The committee agreed that the 12 outcomes should be presented as outcomes related to the pregnant woman (maternal set of outcomes), and outcomes related to the offspring (neonate set of outcomes). The Project Steering Committee agreed unanimously that the outcome selected in the consultation meeting should be included in the final core outcome set and that mother and offspring should have separate outcomes for 'harm'. Therefore, the core outcome set would consist of 13 instead of 12 outcomes. The final core outcome set represents four outcomes related to pregnant women (maternal set): [1] maternal mortality; [2] maternal infection or inflammation; [3] prelabor rupture of membranes; [4] harm to mother from intervention. Nine outcomes related to the offspring (neonate set): [5] gestation age at birth; [6] offspring mortality; [7] birth weight; [8] early neurodevelopmental morbidity; [9] late neurodevelopmental morbidity; [10] gastrointestinal morbidity; [11] infection; [12] respiratory morbidity and [13] harm to offspring from intervention (Box 1).

DISCUSSION

In this project, by utilizing formal consensus methods, we identified a core outcome set of 13 outcomes for comparative health research on preventative interventions for preterm birth in asymptomatic women. These outcomes can be used in future studies, reviews and guidelines on preterm birth prevention.

There are several strengths throughout the different phases of this project. We have followed the guidelines for core outcome set development, as outlined by the Core Outcome Measurement in Effectiveness Trials initiative.³ Second, the method of identification of outcomes was not restricted to the results from a systematic literature review. Questionnaires and interviews were disseminated through conferences and through social media. Third, the parental (patient) perspective was included. This is an important strength as patients can identify outcomes not considered by other stakeholders or within the literature.^{9,10} In this particular core outcome set project we noted that a total of eight outcomes were identified by parental participation that were not identified by prior methods (Appendix 1). Four of these eight outcomes suggested by patients and parents were clustered in the outcome 'late neurodevelopmental morbidity' that was selected in the core outcome set. We hope this will motivate future research to actively involve parents because a recent systematic review concluded that only 16% of reported core outcome studies mentioned that the public has been involved in the process.¹¹ Fourth, we used a Delphi exercise, a well-established method that has the advantage of capturing a large number of geographically distant participants compared to face-to-face discussions. Also, participants have the chance to reconsider their opinion without the pressure to agree with senior or domineering individuals.⁶ This project successfully involved a large number of participants amongst important stakeholder groups and a global representation with participation of middle- and high-income countries. Most of the healthcare professionals involved have a prominent role in their specialties (eg, a high

number of the participants are involved in [inter]national guideline development). This broad involvement of key stakeholders resulted in a core outcome set that should be globally representative and acceptable.

The first limitation of the study is the lack of a formal qualitative analysis of the semistructured interviews with patients and that all patients involved were representing a white ethnic group only. Another limitation is that the stakeholder group representation at the consultation meeting was not reflective of the composition of the group during the Delphi process. Although all stakeholders were represented at the consultation meeting, specifically the midwives, neonatologists and parents (patients), representatives were underrepresented at the consultation meeting. It is possible that the two parents attending the meeting could have found it difficult to argue against the healthcare professionals. The Project Steering Committee addressed this underrepresentation of some stakeholders at the consultation meeting. First, only the outcomes that did not reach full consensus (i.e. consensus 'in' by one or more stakeholder group) were considered in the consultation meeting. Outcomes that were already consensus 'in' after the Delphi exercise were not discussed (ie, 11 of the 13 outcomes in the core outcome set were already agreed through the Delphi exercise). Second, the analysis of the consultation meeting was based on the voting per 'stakeholder group'. This means that every stakeholder group (and not every individual) had the same weight for the decision to include an outcome as a core outcome set. Still, we cannot exclude the possibility that some outcomes were not scored as consensus 'in' due to underrepresentation of some stakeholder groups. A core outcome set is therefore not static and can be adjusted and reviewed in the future.

In the Delphi exercise there were two individuals reporting that they represented the industry as their main stakeholder group. In the analysis we incorporated their outcomes to the second stakeholder group they also belonged to (eg, obstetrician and researcher). Finally, the method of reflecting the first Delphi results to all participants prior to the second Delphi survey might have influenced the second Delphi. Besides the individual stakeholder group responses that were relayed back to the participants by email, we summarised the total group responses in the survey. Because the whole group summary will be affected by the number per stakeholder group, participants may have been influenced by this without realizing that some groups were over- or underrepresented. However, by reflecting both responses (per stakeholder and total responses), we felt that participants were receiving a complete overview of the results.

The proposed 13 core outcomes guide researchers on *what* to measure. It does not tell researchers *how* to measure these outcomes by specifying an outcome measurement instrument and definition for each specific outcome domain. Guidance for selecting high-quality outcome measurement instruments are now being written by the Core Outcome Measurement Instrument Selection project group.¹² In preterm birth, a high quality outcome measurement instrument and definitions are being developed in a separate project.¹³ Until then, we encourage researchers to annotate how an outcome was actually measured and provide the definition used in each trial.

Furthermore, once an outcome set is chosen there may be continued concern that the choices of outcomes in the set do not fit the need of a particular study. Researchers will have their own hypothesis to test, and therefore will need to consider the outcome(s) that reflect their specific hypothesis. Besides collecting hypothesis-specific outcomes, data should be collected and reported on the core outcome set.

Studies on evaluation of treatments in symptomatic woman (like tocolytics) might consider to use this core outcome set in addition to the use of outcomes that are also relevant for that particular study population, for example, 'successful prolongation of pregnancy for 48 hours or longer'. The selection and evaluation of the importance of those particular outcomes are beyond the scope of this core outcome set project.

Consistency of measurements and reporting of the core outcome set in trials is only the first step in the attempt to improve impact and reduce waste.¹⁴ To address possible barriers to the awareness of the core outcome set, a journal editors initiative, Core Outcomes in Women's Health, is encouraging researchers to implement core outcome sets in women's health studies.¹⁵ More than 70 women's health journals are now participating in this initiative (www.crown-initiative.org). Also funders could have an important role encouraging consideration of a core outcome set.¹³

Based on a review of the literature (MEDLINE and EMBASE search 'premature infant [MeSH] AND core outcomes set') and search on the Core Outcome Measures in Effectiveness Trials initiative website (<http://www.comet-initiative.org/studies/search>). This is the first core outcome set developed to ensure consistency in preterm birth prevention research. The initiative from the James Lind Alliance (a partnership regarding research priorities) registered a core outcome sets for very preterm birth from patient perspectives on the Core Outcome Measures in Effectiveness Trials website. This project is still ongoing (www.comet-initiative.org/studies/details/256). It will be of interest to compare the results of these two approaches. Earlier work on a core outcome set for *maternity care* reported 48 outcomes to consider.¹⁶ This core outcome set did not target preterm birth research specifically, and we think that the set of 13 outcomes we recommend here will be more applicable to preterm birth prevention research. The importance of reporting all crucial outcomes has been highlighted in a recent systematic review, which concluded that most published trials in preterm birth are missing information on one of the most crucial outcomes in this population: chronic lung disease.² Although this project does not provide definitions and give advice to which outcome measurement instruments should be used, we would like to suggest that the outcome named as 'chronic lung disease' is captured by the outcome 'respiratory morbidity of the offspring' that is used in this core outcome set project.

In a related project involving the Global Obstetrics Network (www.globalobstetricsnetwork.org), 15 planned trials focusing on the use of vaginal pessary for prevention of preterm birth have already expressed their intention to include this core outcome set in the study protocols and case report forms to facilitate a prospective individual patient data analysis collaboration (see further details above in the method section 'consultation meeting'). The participating research teams have the intention to use the same methods to measure these outcomes and use the same definitions across studies as well.

Even if researchers have the intention to comply with the core outcome set, it is likely that some core outcomes may be difficult to collect. One such example is 'long term neurodevelopment'. This is an outcome that is often not collected due to logistics or lack of funding. In a recent review, only 16% of large obstetrical trials were able to report on follow-up,¹⁷ and only one study used this outcome as a primary outcome.¹⁸ We hope that the development of core outcomes will provide a strong incentive to researchers to argue for adequate funding to perform a follow-up of their planned study. When researchers fail to collect any of the core outcomes, we encourage them to provide an explanation why this outcome was not collected.

We are confident that the development and implementation of a core outcome set will benefit the patients and health care providers by reducing the chance for reporting bias and improving the interpretation of evidence.¹⁴ It will facilitate individual patient data analyses and allowed adequately powered subgroup analyses.

The core outcome set for studies on preterm birth prevention developed with an international multidisciplinary perspective, when implemented in comparative health research, will ensure that data from all trials can be compared and combined.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Seilin Uhm, Juliet Rayment Cecilia Gonzales Marin, and Silke Mader for their help in contacting parents and patient representatives and preparing the information material for parents; Natalie Cooper and Shakila Thangaratinam for her help in the core outcome set methodology; Christina Vinter and Jan Stener Jørgensen for contacting all delegates of the First European Spontaneous Preterm Birth Congress in Svendborg, Denmark, May 2014; and Aoife Ahern, Irene de Graaf, Timo de Haan, Brent Opmeer, Birgit van der Goes, Felipe Castro, and Stephen Kempley for their help in the Project Steering Committee.

References

1. Meher S, Alfirevic Z. Choice of primary outcomes in randomised trials and systematic reviews evaluating interventions for preterm birth prevention: A systematic review. *BJOG*. 2014;1–9.
2. Ioannidis, JPa; Horbar, JD; Ovelman, CM; Brosseau, Y; Thorlund, K; Buus-Frank, ME; , et al. Completeness of main outcomes across randomized trials in entire discipline: survey of chronic lung disease outcomes in preterm infants. *BMJ*. 2015; 350:e72.
3. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials*. 2012; 13:132. [PubMed: 22867278]
4. Kirkham JJ, Boers M, Tugwell P, Clarke M, Williamson PR. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials*. 2013; 14:324. [PubMed: 24103529]
5. WHO. WHO: recommended definitions, terminology and format for statistical tables related to the perinatal period and use of a new certificate for cause of perinatal deaths. Modifications recommended by FIGO as amended October 14, 1976. *Acta Obstet Gynecol Scand*. 1977; 56:247–53. [PubMed: 560099]
6. 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

7. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol.* 2011; 64:395–400. [PubMed: 21194891]
8. Harman NL, Bruce Ia, Callery P, Tierney S, Sharif MO, O'Brien K, et al. MOMENT--Management of Otitis Media with Effusion in Cleft Palate: protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials.* 2013; 14:70. [PubMed: 23497540]
9. De Wit M, Abma T, Koelewijn-van Loon M, Collins S, Kirwan J. Involving patient research partners has a significant impact on outcomes research: a responsive evaluation of the international OMERACT conferences. *BMJ Open.* 2013; 3:1–12.
10. Hewlett S, De Wit M, Richards P, Quest E, Hughes R, Heiberg T, et al. Patients and professionals as research partners: Challenges, practicalities, and benefits. *Arthritis Care Res.* 2006; 55:676–80.
11. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, et al. Choosing important health³ 13–16 outcomes for comparative effectiveness research: A systematic review. *PLoS One.* 2014; 9(6)
12. Prinsen CC, Vohra S, Rose MR, King-Jones S, Ishaque S, Bhaloo Z, et al. Core Outcome Measures in Effectiveness Trials (COMET) initiative: protocol for an international Delphi study to achieve consensus on how to select outcome measurement instruments for outcomes included in a “core outcome set”. *Trials.* 2014; 15:247. [PubMed: 24962012]
13. Lawn JE, Gravett MG, Nunes TM, Rubens CE, Stanton CS and the GAPPS Review Group. Global report on preterm birth and stillbirth (1 of 7): definitions, description of the burden and opportunities to improve data. *BMC Preg and Chil.* 2010; 10:S1.
14. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet.* 2009; 374:786. [PubMed: 19733780]
15. Khan K. The CROWN Initiative: Journal editors invite researchers to develop core outcomes in women's health. *BJOG.* 2014; 126:201–2.
16. Devane D, Begley CM, Clarke M, Horey D, OBoyle C. Evaluating Maternity Care: A Core Set of Outcomes Measures. *BIRTH.* 2007; 34:164–72. [PubMed: 17542821]
17. Teune MJ, van Wassenaer aG, Malin GL, et al. Long-term child follow-up after large obstetric randomised controlled trials for the evaluation of perinatal interventions: a systematic review of the literature. *BJOG.* 2013; 120:15–22. [PubMed: 23078194]
18. Lees CC, Marlow N, van Wassenaer-Leemhuis A, Arabin B, Bilardo CM, Brezinka C, et al. 2 year neurodevelopmental and intermediate perinatal outcomes in infants with very preterm fetal growth restriction (TRUFFLE): a randomised trial. *Lancet.* 2015; 4:S0140–6736.

Box 1. Final Core Outcome Set of 13 Outcomes Presented as a Maternal and Neonate Set

MATERNAL SET OF OUTCOMES	NEONATAL SET OF OUTCOMES
Maternal mortality	Offspring mortality
Maternal infection or inflammation	Offspring infection
Prelabor rupture of membranes	Gestational age at birth
Harm to mother from intervention	Harm to offspring from intervention
	Birth weight
	Early neurodevelopmental morbidity
	Late neurodevelopmental morbidity
	Gastrointestinal morbidity
	Respiratory morbidity

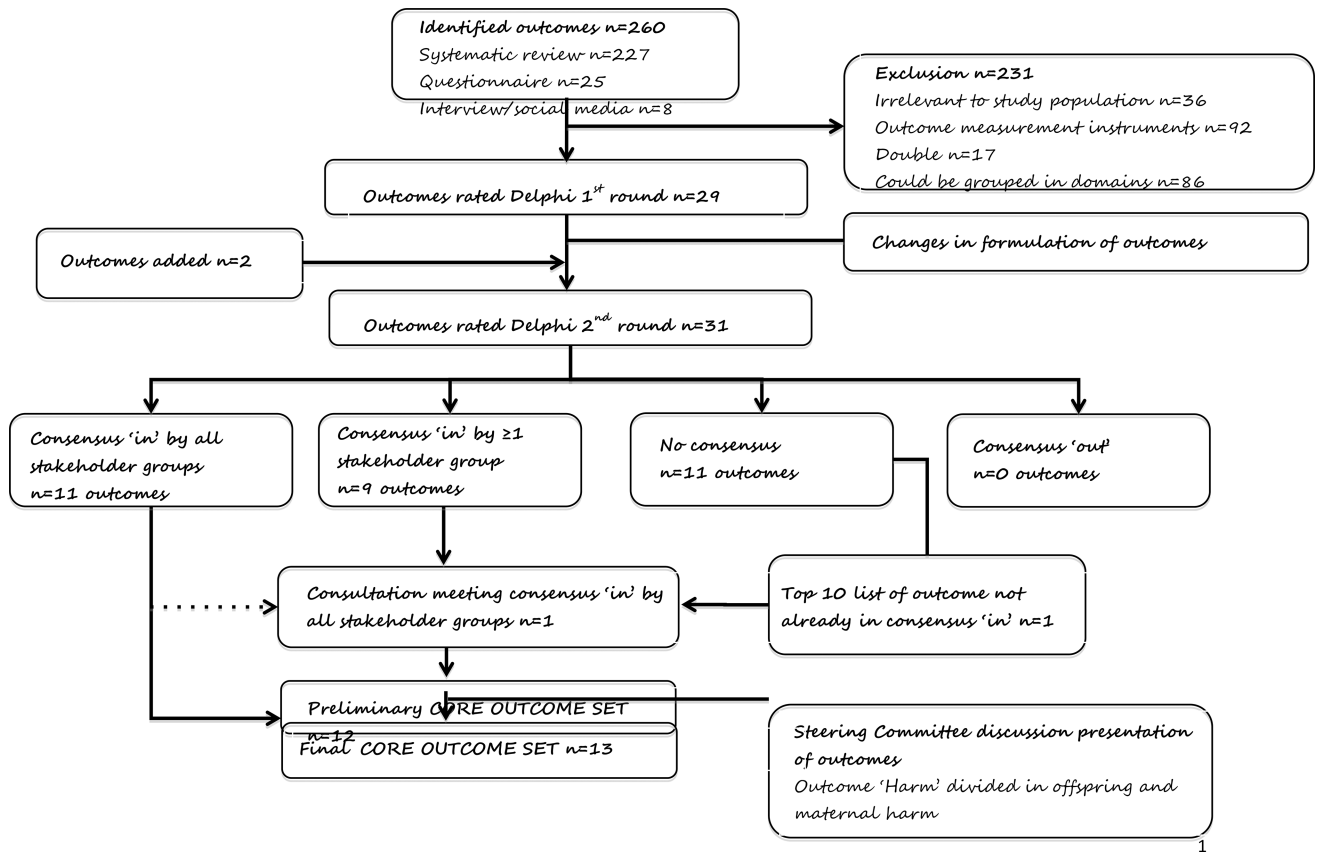


Figure 1.
Flowchart of identification and selection of outcomes.

Table 1

Total number and baseline characteristics of participants of the Delphi survey and consultation meeting.

Stakeholder groups	1st Delphi round (n=195)	2nd Delphi round (n=174)	Consultation meeting (n=29)
Parents n (% of total group)	32 (16)	25 (14)	2 (7)
<i>Response %</i>	<i>84</i>	<i>78</i>	
Midwives n (%)	28 (14)	25 (14)	1 (3)
<i>Response %</i>	<i>78</i>	<i>89</i>	
Neonatologists n (%)	34 (18)	34 (20)	2 (7)
<i>Response %</i>	<i>80</i>	<i>100</i>	
Obstetricians n (%)	62 (32)	55 (32)	14 (48)
<i>Response %</i>	<i>90</i>	<i>89</i>	
Researchers n (%)	39 (20)	35 (20)	10 (34)
<i>Response %</i>	<i>91</i>	<i>90</i>	
<i>Total group response %</i>	<i>86</i>	<i>89</i>	<i>100</i>
Characteristics health professionals			
Main work clinical related %	62	60	57
Main work research related %	36	40	43
Other	2	0	0
Representing other stakeholder groups			
Parent experiencing preterm birth %	1	1	17
Midwife %	6	6	7
Obstetrician %	14	16	14
Neonatologist %	4	5	3
Researcher %	38	39	34
Industry %	2	2	0
Part of CROWN or representing journal %	22	22	18
Part of Cochrane collaboration %	25	27	21
Systematic review related to preterm birth? %	54	54	-
Role in development of national/international guidelines %	60	61	64
Role in allocation of healthcare budgets %	9	9	21
Countries represented n (countries healthcare professionals working in) * †	25	25	8
High-income countries n	16	16	6
Upper-middle-income countries n	6	6	2
Lower-middle-income countries n	3	3	0
Participants middle-income countries n (%)	20 (12)	19 (13)	2 (7)
Characteristics parents			
Female %	91	88	-
Experienced preterm birth <37 weeks %	94	96	-
Once %	69	72	-

Stateholder groups	1st Delphi round (n=195)	2nd Delphi round (n=174)	Consultation meeting (n=29)
Twice %	25	24	-
Gestational age most premature baby median (range)	30 (24–35)	30 (24–35)	-
Highest degree of education median (range)	Master's degree (high school to doctorate degree)	Master's degree (high school to doctorate degree)	-
Ethnic group white %	100	100	-
Involved in research before %	59	60	-
Participated in study %	31	36	-
Helped in a study giving advice from parental/patient perspective %	9	4	-
Worked as a researcher %	19	20	-
Represented countries of residence n [‡]	6	6	-
High-income countries n	5	5	-
Upper-middle-income countries n	1	1	-

* Represented countries healthcare professionals: Argentina, Australia, Brazil, Canada, Chile, China, Denmark, Egypt, Germany, Hong Kong, Iran, Ireland, Italy, Lebanon, Mexico, Nigeria, Pakistan, Qatar, South Africa, Spain, Switzerland, The Netherlands, United Kingdom, Uruguay, USA.

[‡] Represented countries consultation meeting: Australia, Brazil, China, France, Spain, The Netherlands, UK, USA

[‡] Represented countries parents: Greece, Ireland, Serbia, The Netherlands, United Kingdom, USA.

Table 2

Delphi second round: top 10 outcomes for each stakeholder Group.

Total group (n=174)	Parents (n=25)	Midwives (n=25)	Obstetricians (n=55)	Neonatologists (n=34)	Researchers (n=35)
Gestational age at delivery (93%)	Maternal mortality (76%)	Gestational age at delivery (92%)	Gestational age at delivery (100%)	Gestational age at delivery (91%)	Gestational age at delivery (97%)
Offspring mortality (81%)	Gestational age at delivery (72%)	Offspring mortality (76%)	Offspring mortality (89%)	Late neurodevelopmental morbidity (79%)	Offspring mortality (85%)
Early neurodevelopmental morbidity (72%)	Offspring mortality (72%)	Maternal infection (64%)	Early neurodevelopmental morbidity (85%)	Offspring mortality (71%)	Early neurodevelopmental morbidity (82%)
Late neurodevelopmental morbidity (72%)	Early neurodevelopmental morbidity (68%)	Respiratory morbidity (60%)	Late neurodevelopmental morbidity (80%)	Early neurodevelopmental morbidity (68%)	Respiratory morbidity (74%)
Respiratory morbidity (67%)	Offspring infection (68%)	Offspring infection (56%)	Respiratory morbidity (80%)	Maternal mortality (59%)	Maternal mortality (62%)
Maternal mortality (60%)	Late neurodevelopmental morbidity (60%)	Prelabor rupture of membranes (56%)	Maternal mortality (59%)	Birthweight (56%)*	Late neurodevelopmental morbidity (56%)
Maternal infection (55%)	Respiratory morbidity (56%)	Late neurodevelopmental morbidity (52%)	Maternal infection (57%)	Respiratory morbidity (53%)	Mode of delivery (53%) [†]
Offspring infection (52%)	Maternal infection (56%)	Maternal mortality (48%)	Offspring infection (57%)	Maternal infection (53%)	Harm (53%)
Birthweight (50%)*	Gastro-intestinal morbidity (44%)	Harm (48%)	Birthweight (57%)*	Health service utilization offspring (41%)*	Birthweight (50%)*
Prelabor rupture of membranes (43%)	Circulatory morbidity (44%)*	Early neurodevelopmental morbidity (44%)	Mode of delivery (56%) [†]	Mode of delivery (38%) [†]	Prelabor rupture of membranes (50%)

All outcomes are also in the consensus 'in' list of all stakeholder groups, except for:

* Outcome also in consensus 'in' list by one or more stakeholder groups.

[†] Outcome in 'no' consensus list by all stakeholder groups.