



Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process

KJ Bunch,^a B Allin,^a M Jolly,^b T Hardie,^b M Knight^a

^a Policy Research Unit in Maternal Health and Care, National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK ^b NHS England, London, UK

Correspondence: KJ Bunch, Policy Research Unit in Maternal Health and Care, National Perinatal Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Old Road Campus, Headington, Oxford OX3 7LF, UK. Email: kathryn.bunch@npeu.ox.ac.uk

Accepted 13 April 2018. Published Online 14 June 2018.

Objective To develop a core metric set to monitor the quality of maternity care.

Design Delphi process followed by a face-to-face consensus meeting.

Setting English maternity units.

Population Three representative expert panels: service designers, providers and users.

Main outcome measures Maternity care metrics judged important by participants.

Methods Participants were asked to complete a two-phase Delphi process, scoring metrics from existing local maternity dashboards. A consensus meeting discussed the results and re-scored the metrics.

Results In all, 125 distinct metrics across six domains were identified from existing dashboards. Following the consensus meeting, 14 metrics met the inclusion criteria for the final core set: smoking rate at booking; rate of birth without intervention; caesarean section delivery rate in Robson group 1 women; caesarean section delivery rate in Robson group 2 women; caesarean section delivery rate in Robson group 5 women; third-

and fourth-degree tear rate among women delivering vaginally; rate of postpartum haemorrhage of ≥ 1500 ml; rate of successful vaginal birth after a single previous caesarean section; smoking rate at delivery; proportion of babies born at term with an Apgar score < 7 at 5 minutes; proportion of babies born at term admitted to the neonatal intensive care unit; proportion of babies readmitted to hospital at < 30 days of age; breastfeeding initiation rate; and breastfeeding rate at 6–8 weeks.

Conclusions Core outcome set methodology can be used to incorporate the views of key stakeholders in developing a core metric set to monitor the quality of care in maternity units, thus enabling improvement.

Keywords Core outcome set, dashboard, Delphi process, indicator, maternity service, quality improvement, quality of care.

Tweetable abstract Achieving consensus on core metrics for monitoring the quality of maternity care.

Linked article This article is commented on by BD Einerson, p. 1619 in this issue. To view this mini commentary visit <https://doi.org/10.1111/1471-0528.15332>.

Please cite this paper as: Bunch KJ, Allin B, Jolly M, Hardie T, Knight M. Developing a set of consensus indicators to support maternity service quality improvement: using Core Outcome Set methodology including a Delphi process. *BJOG* 2018;125:1612–1618.

Introduction

There is an increasing focus internationally on improving the quality of maternity care and the development and application of measures to drive quality improvement.^{1–3} Two key policy drivers relevant to improving the quality of maternity care have recently been introduced in England. The first of these was the announcement by the government in November 2015 of a national ambition to reduce maternal mortality, stillbirth, neonatal mortality and

serious neonatal injury by 20% by 2020, and 50% by 2030.⁴ The second is the development of a maternity transformation programme to implement recommendations of the ‘Better Births’ report,⁵ which aims to improve the quality of care for women, babies and their families.

Achieving quality improvement requires local providers to have better information about the quality of their services. In response, many maternity units or networks have developed, or are developing, sets of service quality indicators, referred to as maternity ‘dashboards’. These are

monitored on a monthly basis using information obtained from routine hospital data.

Dashboards currently in use are variable, with some elements included that are responsive to local needs, and other 'core' metrics that could be generally applicable. The methods used by different units to select metrics are often unclear. Not all chosen metrics are based on evidence of utility or responsiveness to change, nor are the required data always routinely available.

Other authors have noted the challenges of developing metrics to drive quality improvement, and noted that gaps in the evidence base are a problem.¹ Over the past decade, researchers have developed methodology for identifying outcomes that key stakeholders deem important in defining successful treatment of specific conditions^{6–11} with much of this work taking place within the COMET initiative. Such methodology generally consists of four stages: a systematic review to identify reported outcomes, qualitative work to identify outcomes considered important by patients, a Delphi process to prioritise outcomes, and finally a face-to-face consensus meeting. The resulting 'Core Outcome Set' can be used in all future studies comparing safety and efficacy of treatments.

The overall objective of this study was to investigate whether it was possible to adapt 'Core Outcome Set' development methodology to identify in a timely, transparent and robust manner, quality of care measures that could have genuine utility in improving patient experience and outcomes across English maternity units. Although we describe work carried out in England, the techniques used could equally well be applied in other settings.

Methods

Identification of candidate metrics

In the absence of previous research studies, existing dashboards were reviewed by the National Child and Maternal Health Intelligence Network (ChiMat) and NHS England staff to produce a comprehensive list of metrics used in current dashboards in England. Duplicate metrics were removed, and the remaining unique metrics were grouped into six domains: antenatal, maternal, neonatal, mental health, public health and workforce-related and taken forward for assessment in the first phase of a Delphi process.

Panel formation and recruitment

To ensure that the views of all key stakeholder groups were represented within the core indicator set, experts were recruited from across the breadth of maternity care. However, to facilitate feedback of meaningful data in an easily interpreted way during the Delphi process, stakeholder groups were combined into three panels:

- Service design panel – individuals whose responsibilities included commissioning services, maternity service policy, population health services, or national audit and research, i.e. responsibility for healthcare/service improvement at a population level.
- Service provision panel – clinicians/managers whose responsibility included direct provision of maternity services, i.e. responsibility for individual-level health care.
- Public panel – users and representatives of charities and other voluntary organisations working in the maternity arena.

Recruitment was conducted according to an adaptation of methodology described by Okoli and Pawlowski.¹² Members of the project management group populated each category of stakeholder with names of experts known to them, including representatives from all maternity networks across England, and significant third sector organisations working in the area of maternity care. Strategies to identify further experts in each category were then developed, including contacting lead individuals in each maternity network to request further nominations.

Each identified expert was sent an information pack by email. The information packs explained that they had been identified as having expertise in the area of maternity care and quality improvement and asked whether they would consider participating in the development of a core indicator set. Each participant's invitation contained a link to the data collection website through which the Delphi process was conducted. Participants were asked to confirm their participation by following the link, at which point they were able to proceed immediately to phase 1 of the Delphi process.

Delphi process phase 1

Data collection

Participants were presented with the list of candidate metrics, and asked to score each from 1 to 9 based upon their importance in monitoring the quality of maternity care. The GRADE scale of measurement was chosen for use in scoring metrics, based on recommendations from the COMET initiative.¹¹ Participants were also offered the opportunity to comment on the metrics and list any additional metrics they considered important that had not been assessed in phase 1. Participants were sent up to three reminders to complete the phase. Participants who had not completed the questionnaire within 4 weeks were deemed not to have completed phase 1 and were excluded from phase 2.

Analysis

Scores were analysed separately for each panel, with descriptive statistics calculated. All metrics were carried

forward to phase 2. Two reviewers (KJB and BA) independently assessed additional metrics suggested by phase 1 participants to determine if they represented *de novo* metrics not already listed. Uncertainties were resolved by a third reviewer (MK) and the final list of additional metrics was reviewed by the project management group. *De novo* metrics listed by at least one expert were taken forward to phase 2 of the Delphi process.

Delphi Process Phase 2

Data collection

Experts completing phase 1 were invited to participate in phase 2, and asked to re-score each metric based on:

- the phase 1 score they had assigned it
- graphical and numerical representations of their panel's scores for that metric from phase 1.

Again up to three reminders were sent to participants who had not completed the second questionnaire.

Analysis

Metric scores were analysed separately for each panel, with descriptive statistics calculated. Metrics were then grouped according to whether they reached 'consensus in' in each panel according to the recommendations of the COMET initiative.¹¹ Metrics were considered 'consensus in' where:

- $\geq 70\%$ of participants rated the metric 7–9 (high importance) and
- $< 15\%$ rated it as 1–3 (low importance).

Generation of core metric set – consensus meeting

Metrics determined as 'consensus in' by at least two of the three panels were carried forward to the consensus meeting. In addition, consensus group members had an opportunity to promote metrics from the list of those that did not reach consensus. The consensus meeting included representative experts who had participated in the Delphi process, together with key organisational stakeholders with relevant expertise in data sources and metric measurement.

Consensus group members were asked to consider the following principles when reviewing metrics:

- The metrics should be important to drive clinical quality improvement.
- They should be measureable using current data sources (aspirational metrics were recorded for future data set development).
- They should be useful when monitored on a monthly basis, informed by rarity of the event. As a guide, events with a frequency of $< 1\%$ were considered unsuitable for this purpose and more suitable for monitoring annually through existing audits such as the Mothers and Babies: Reducing Risk through Audits and Confidential Enquiries

(MBRRACE-UK) surveillance of stillbirths, neonatal and maternal deaths^{13,14} or the National Maternity and Perinatal Audit.¹⁵

- Where metrics were similar, scoring should reflect prioritisation of those considered most important.
- The final dashboard tool would have a configurable element such that metrics could be viewed according to subgroups, for example according to gestation at birth, mode of delivery, place of birth, twin or triplet pregnancies. Therefore subgroups of a more general metric should not be included.
- Metrics should involve rates and not simply numbers of events. Numerator and denominator information underlying each metric (for example the total number of births) would be available to view by units and hence would provide the overall contextual information about, for example, unit size.
- Poor data quality should not be a reason to choose to score the metric low, as inclusion of a metric in a dashboard may drive improvement in data quality.

Open, group discussion was held about each metric to explore the reasons given both for inclusion and exclusion by different categories of experts. Following discussion, participants were asked to re-score each metric as per the Delphi process. To remove the pressure that dominant personalities can exert on decision-making, voting was electronic and anonymous. The chair remained neutral and did not vote. Metrics that achieved 'consensus in' status following scoring at the consensus meeting were included in the core metric set.

Results

In all, 125 distinct metrics were identified from existing maternity dashboards and were listed under the six domain headings for phase 1 of the Delphi process. A further 19 *de novo* metrics were nominated in phase 1 and added at phase 2, which therefore included 144 metrics. An additional domain was also added at this stage to explore participants' views on the importance of being able to view indicators according to parity, plurality, gestational age and Robson group.¹⁶ A full listing of the metrics scored during phases 1 and 2 of the Delphi Process is given in Table S1.

The numbers of participants within each panel and the numbers completing the different stages of the Delphi process are shown in Table 1 and further details of the composition of the panels is shown in Table S2.

Following scoring in phase two, 33 metrics met the criteria for consideration at the consensus meeting across all three panels and a further 46 metrics met the criteria across two of the three panels. Seventy-nine metrics were therefore taken forward for discussion at the consensus meeting.

Table 1. Numbers of experts participating in the Delphi process and Consensus meeting

Panel	Service design	Service provision	Public	Total
Delphi process				
Invited	26	64	11	101
Completed phase 1	22 (85%)	53 (83%)	7 (64%)	82* (81%)
Completed phase 2	21 (81%)	44 (69%)	7 (64%)	72 (71%)
Consensus meeting	13	6	0	19

*Of the 19 phase 1 non-responders, two declined and 17 had not responded after three reminders.

Nineteen participants voted at the consensus meeting, including 13 service designers of whom nine had completed both phases of the Delphi process and six service providers of whom five had completed both Delphi phases.

Fourteen metrics met the criteria for inclusion in the final dashboard set:

- smoking rate at booking
- rate of birth without intervention
- caesarean section delivery rate in Robson group 1 women
- caesarean section delivery rate in Robson group 2 women
- caesarean section delivery rate in Robson group 5 women
- third- and fourth-degree tear rate among women delivering vaginally
- rate of postpartum haemorrhage of ≥ 1500 ml
- rate of successful vaginal birth after a single previous caesarean section
- smoking rate at delivery
- proportion of babies born at term with an Apgar score < 7 at 5 minutes
- proportion of babies born at term admitted to the neonatal intensive care unit
- proportion of babies readmitted to hospital at < 30 days of age
- breastfeeding initiation rate
- breastfeeding rate at 6–8 weeks

Details of the scoring for these metrics during phases 1 and 2 of the Delphi process and at the consensus meeting are listed in Table 2.

One of these 14 metrics, Birth without intervention (vaginal birth without induction, epidural, augmentation, forceps, ventouse or episiotomy), was not being used in any of the existing dashboards, but the consensus group argued for its inclusion on the basis that the consensus list lacked a useful metric concerning non-intervention. A further included metric, Apgar score < 7 at 5 minutes in term

babies, was promoted from the list of metrics reaching only one panel consensus. The meeting was influenced in this decision by the fact that a parallel metric, Apgar score < 4 at 5 minutes in term babies, had been considered important by all three panels. However, it was agreed that such low Apgar scores would arise too rarely to be useful for monitoring on a monthly basis.

Discussion

Main findings

This study successfully used 'Core Outcome Set' development methodology to identify a set of 14 core metrics that key stakeholders deemed important in assessing aspects of maternity care amenable to quality improvement initiatives. These metrics can all be assessed using routinely collected hospital data, and so provide a practical system for rapidly delivering meaningful feedback to trusts in relation to their clinical performance. The metrics identified span the breadth of maternity care, including public health, maternal and neonatal outcomes.

Strengths and limitations

Use of the structured core outcome set development methodology¹¹ allowed for creation of a core indicator set that is relevant to clinical care over a short, 6-month time-period, at minimal cost, using robust, transparent methodology. Specifically, use of the Delphi process allowed for wide participation from staff throughout maternity services, commissioning, policy and public health organisations as well as other key stakeholders including service user representatives, contributors to national audits and researchers. Discussion at the consensus meeting then allowed for additional, in-depth consideration of important aspects such as practicality of the measures and configuration of other dashboard features including the use of sub-groups, which will enhance the potential of the core metrics for quality improvement among specific risk groups.

Despite the large number of participants in the study, a key limitation is the relatively small size of the public panel. It was considered important by the dashboard developers that participants had not only expertise in identification of metrics, but also awareness of which metrics could be reliably reported from existing routinely collected data. This was to ensure that the metrics identified were deliverable; however, there were few public representatives with the requisite expertise and as a result recruitment to this panel was difficult. Those who participated were solely from third sector organisations, and it is possible that the inclusion of maternity service users may have contributed different opinions concerning important metrics.¹⁷

A large number of metrics met consensus after the initial stages of the Delphi process, the majority of which

Table 2. Scoring details for final consensus maternity dashboard metrics

Metric	Phase 1 % scoring 7–9 by panel			Phase 2 % scoring 7–9 by panel			Consensus meeting % scoring 7–9
	Service design	Service provision	Public	Service design	Service provision	Public	
Smoking rate at booking	73	54	57	81	71	71	94
Rate of birth without intervention	–	–	–	–	–	–	100*
CS rate in Robson group 1 women	83	73	43	95	88	33	100
CS rate in Robson group 2 women	–	–	–	–	–	–	100**
CS rate in Robson group 5 women	–	–	–	–	–	–	100**
Third- and fourth-degree tear rate, non-instrumental births	80	85	50	90	89	50	88***
Third- and fourth-degree tear rate, assisted births	84	83	33	90	96	17	
PPH rate \geq 1500 ml	83	60	86	81	68	100	100
Successful VBAC rate	61	69	67	73	72	33	87
Smoking rate at delivery	77	60	57	86	68	71	100
Apgar <7 5 min at term rate	70	71	83	86	90	100	89****
NICU admission rate at term	73	78	86	95	93	100	94
Neonatal readmission <30 days rate	68	54	71	81	61	86	89
Breastfeeding initiation rate	59	67	67	71	73	83	78
Breastfeeding rate at 6–8 weeks	50	49	67	76	52	100	100

CS, caesarean section; NICU, neonatal intensive care unit; PPH, postpartum haemorrhage; VBAC, vaginal birth after caesarean section.

*Not scored in phases 1 and 2.

**Not scored in phases 1 and 2, but consensus group considered caesarean section rate should be monitored for Robson groups 1, 2 and 5.

***Phases 1 and 2 asked panellists to score third- and fourth-degree tear separately for non-instrumental vaginal births and assisted births, consensus meeting decided to combine groups.

****Panellists were asked to score Apgar <4 at 5 minutes in phases 1 and 2, consensus meeting decided that Apgar <7 was more appropriate.

represented important outcomes in maternity care, but which were not, following discussion at the consensus meeting, considered useful for monitoring on a monthly basis to improve the quality of clinical care.

Although application of the Delphi process ensures objectivity in determining which of the candidate metrics are included in the core set, the initial identification of the candidate metrics could not be informed by a systematic review, as dashboard work is typically not published in traditional formats. Where core outcome sets are defined based on published clinical trials, a systematic review of such trials provides an objective method of identifying candidate outcomes. There is, however, much less consistency in the reporting of maternal, neonatal and perinatal outcomes by hospitals with the result that metrics were included in the initial list only if a hospital had reported their use to NHS England. Giving participants the opportunity to suggest further important metrics during phase 1 of the Delphi process should have prevented the omission of any essential metrics, but further investigation of how to identify candidate metrics is needed.

Interpretation

The remit of our study was to identify metrics for monthly monitoring within England, whereas other studies have had broader or different focuses. Devane et al.¹⁰ sought to identify measures to evaluate maternity care internationally and produced a much longer set of outcomes. Nevertheless, there is considerable overlap between the outcomes identified by their study and ours, particularly in the areas of mode of delivery, postpartum haemorrhage, Apgar score and breastfeeding uptake. A more recent report by Iriye et al.² focuses on quality measures for high-risk pregnancies in the USA and the recommendations considered are almost exclusively clinical. This is in contrast to our attempt to identify measures applicable across the risk spectrum and to encompass public health perspectives as well as those of service users.

As our study was designed to assist with the construction of a dashboard for monthly monitoring, outcomes with a frequency of <1% were unsuitable for inclusion. Hence several major adverse outcomes, most notably perinatal death, do not appear in the final core metrics. These are more

appropriately monitored nationally. In the UK the MBRRACE-UK programme uses national data to identify system-level actions for improvement and this additionally helps to reduce the quantity of costly litigation arising in this area.

A key element in our study was the consensus meeting, which opened with a discussion of the principles of identifying useful metrics and these principles clearly guided much of the subsequent voting, reflected in very consistent scores across all participants. This highlighted the importance of this final face-to-face meeting phase to clearly establish the principles by which each participant was making an assessment. The additional value of the consensus meeting was the presence of experts who could advise in detail on the use of routine sources of data, to ensure that the metrics identified were immediately practical to monitor robustly using data already collected. The consensus meeting further benefitted from inclusion of experts with responsibility for future iterations of the mandated national maternity data collection, allowing for aspirational metrics to be discussed.

The consensus group highlighted particularly that there are currently no suitable metrics evaluating user experience of care and clearly, when these are developed, the service user and third sector organisation perspective will be essential. In future studies, significant effort will have to be put into ensuring that there is a mechanism in place so that dashboard developers can ensure the expertise of participants is met, while also ensuring that the patient and service user voice is heard strongly throughout all stages of the core indicator set development process.

Conclusion

This study has incorporated the views of key stakeholders to develop a core metric set that can be used for monitoring the quality of care provided by maternity units in England. Standardisation of data for primary use collected by different units will improve data used for secondary purposes and aid comparison of hospitals, and also ensure that hospitals are assessed against outcomes that are useful in clinical decision-making. In the longer term, this will allow hospitals to assess where improvements are required, and implement changes to care provision that will be beneficial to women, children and individual hospitals. This study has shown that it is possible to apply robust, transparent methodology to selection of metrics that are used for assessing the quality of care provided in the NHS. In an era where there is significant pressure from policy organisations to demonstrate high-quality, effective care, yet there is little thought given to how that care should be assessed, data fed back, or care improved, it is essential that an evidence-based approach is developed both for identification

of hospital-level metrics, but also, where mandated, surgeon- or physician-specific metrics. We believe that the approach used here is a way forward. Further research is required to test whether and in what ways monitoring of these metrics drives change.

Acknowledgements

The authors would like to acknowledge helpful comments from Dominic Gair and Paula Curnow (NHS Digital) and Hannah Knight (National Maternity and Perinatal Audit) on the data fields/methodology for deriving the final metrics. Essential contributions were made by the Project Management Group, which, in addition to the authors, included Helen Duncan and Helen Smith (Public Health England) together with Thelma Goddard and Jennifer Stanley (NHS England). We are also grateful to all those who participated in the Delphi process and the final consensus meeting.

Disclosure of interests

None declared. Completed disclosure of interests form available to view online as supporting information.

Contribution to authorship

MK and MJ conceived the study. MK, MJ, KJB, BA and TH contributed to study design. KJB and BA designed and implemented the Delphi Process and analysed the resulting data. KJB wrote the first draft of the manuscript. All authors interpreted the results and edited the manuscript.

Details of ethics approval

No ethical approval was required for this study.

Funding

This paper reports an independent study that is funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.

Supporting Information

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

Table S1. Candidate metrics scored in phases 1 and 2 of the Delphi Process.

Table S2. Experts invited to join the three panels. ■

References

- 1 Adirim T, Meade K, Mistry K. A new era in quality measurement: the development and application of quality measures. *Pediatrics* 2017;139:e20163442.

- 2 Iriye BK, Gregory KD, Saade GR, Grobman WA, Brown HL. Quality measures in high-risk pregnancies: executive Summary of a Cooperative Workshop of the Society for Maternal-Fetal Medicine, National Institute of Child Health and Human Development, and the American College of Obstetricians and Gynecologists. *Am J Obstet Gynecol* 2017;217:B2–25.
- 3 Kannan V, Fish JS, Mutz JM, Carrington AR, Lai K, Davis LS, et al. Rapid development of specialty population registries and quality measures from electronic health record data. An Agile Framework. *Meth Inform Med* 2017;56:e74–83.
- 4 Department of Health. New ambition to halve rate of stillbirths and infant deaths 2015 [www.gov.uk/government/news/new-ambition-to-halve-rate-of-stillbirths-and-infant-deaths]. Accessed 15 March 2017.
- 5 The National Maternity Review. *Better Births*. London: The National Maternity Review; 2015.
- 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:e1000393.
- 7 Duffy JMN, Rolph R, Gale C, Hirsch M, Khan KS, Ziebland S, et al. Core outcome sets in women's and newborn health: a systematic review. *BJOG* 2017;124:1481–9.
- 8 Egan AM, Galjaard S, Maresh MJA, Loeken MR, Napoli A, Anastasiou E, et al. A core outcome set for studies evaluating the effectiveness of prepregnancy care for women with pregestational diabetes. *Diabetologia* 2017;60:1190–6.
- 9 Schaap T, Bloemenkamp K, Deneux-Tharoux C, Knight M, Langhoff-Roos J, Sullivan E, et al. Defining definitions: a Delphi study to develop a core outcome set for conditions of severe maternal morbidity. *BJOG* 2017; <https://doi.org/10.1111/1471-0528.14833>.
- 10 Devane D, Begley CM, Clarke M, Horey D, Oboyle C. Evaluating maternity care: a core set of outcome measures. *Birth* 2007;34:164–72.
- 11 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.
- 12 Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manag* 2004;42:15–29.
- 13 Knight MNM, Tuffnell D, Kenyon S, Shakespeare J, Gray R, Kurinczuk JJ editors, on behalf of MBRRACE-UK. *Saving Lives, Improving Mothers' Care – Surveillance of Maternal Deaths in the UK 2011–13 and Lessons Learned to Inform Maternity Care from the UK and Ireland Confidential Enquiries into Maternal Deaths and Morbidity 2009–13*. Oxford: National Perinatal Epidemiology Unit, University of Oxford; 2015.
- 14 Manktelow BM, Smith LK, Seaton SA, Hyman-Taylor P, Kurinczuk JJ, Field DJ, et al. *Perinatal Mortality Surveillance Report UK Perinatal Deaths for births from January to December 2014*. Leicester: The Infant Mortality and Morbidity Group, Department of Health Sciences, University of Leicester; 2016.
- 15 National Maternity and Perinatal Audit (NMPA). 2016 [<http://www.maternityaudit.org.uk>]. Accessed 25 January 2018.
- 16 Robson M, Murphy M, Byrne F. Quality assurance: the 10-Group Classification System (Robson classification), induction of labor, and cesarean delivery. *Int J Gynaecol Obstet* 2015;131(Suppl 1):S23–7.
- 17 Harman NL, Bruce IA, Kirkham JJ, Tierney S, Callery P, O'Brien K, et al. The importance of integration of stakeholder views in core outcome set development: otitis media with effusion in children with cleft palate. *PLoS ONE* 2015;10:e0129514.