Essential variables for reporting research studies on fetal growth restriction

Fetal Growth Restriction Minimum Reporting Set Working Group.; Khalil, Asma; Gordijn, Sanne J; Beune, Irene M; Wynia, Klaske; Ganzevoort, Wessel; Figueras, Francesc; Kingdom, John; Marlow, Neil; Papageorghiou, Aris T

Published in:
Ultrasound in Obstetrics & Gynecology

DOI:
10.1002/uog.19196

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Final author's version (accepted by publisher, after peer review)

Publication date:
2019

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Essential variables for reporting research studies on fetal growth restriction – a Delphi consensus

Asma Khalil¹,², Sanne J Gordijn³, Irene M. Beune³, Klaske Wynia⁴, Wessel Ganzevoort⁵, Francesc Figueras⁶, John Kingdom⁷, Neil Marlow⁸, Aris T Papageorghiou¹, Neil Sebire⁹, Jennifer Zeitlin¹⁰, Ahmet A Baschat¹¹.

(Fetal Growth Restriction Minimum Reporting Set Working Group).

¹Fetal Medicine Unit, St George’s Hospital, St George’s University of London, UK. Cranmer Terrace, London SW17 0RE
²Vascular Biology Research Centre, Molecular and Clinical Sciences Research Institute, St George’s University of London, Cranmer Terrace, London SW17 0RE, UK.
³Department of Obstetrics and Gynecology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands.
⁴Department of Health Sciences, Community and Occupational Medicine, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
⁵Department of Obstetrics and Gynecology, Academic Medical Center Amsterdam, University of Amsterdam, Amsterdam, Netherlands
⁶Hospital Clinic de Barcelona, Spain
⁷University of Toronto, Canada
⁸Institute for Women’s Health, University College London
⁹Institute of Child Health, University College London
¹⁰Center for Epidemiology and Biostatistics, Paris, France
¹¹Center for Fetal Therapy, Department of Gynecology & Obstetrics, Johns Hopkins School of Medicine, Baltimore, MD, USA

Corresponding author:

Professor Asma Khalil
Fetal Maternal Medicine Unit
St George’s University of London
London SW17 0RE
E-mail: akhalil@sgul.ac.uk; asmakhalli79@googlemail.com

Short Title: Reporting research studies on growth restriction

Keywords: growth restriction, Delphi consensus, research, studies, variables

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/uog.19196
ABSTRACT

Objective: To achieve consensus on the minimum reporting set of study variables for fetal growth restriction (FGR) research studies. Determination of a list of variables considered essential to be reported independent of a specific hypothesis is likely to improve the study quality by inclusion of essential end-points, enhancing the consistency among studies and minimizing potential confounding. This in turn can accelerate generation of generalizable knowledge either by standardization of individual study designs or by enhancing the possibility of individual patient data meta-analysis merging a number of uniformly structured datasets.

Study Design: An expert panel, identified based on their publication record as lead or senior authors in studies of FGR, was requested to select a set of essential study parameters from a literature-based list utilizing the Delphi consensus methodology. We collected responses in four consecutive rounds by online questionnaires presented to panelists through a unique token-secured link for each round. Variables were selected in three rounds based on the concurrence on Likert scale scoring. In the final round, retained parameters were categorized as essential (to be reported in all FGR studies) or recommended (important but not mandatory).

Results: 87/100 experts agreed to participate and 62 (71%) completed all four rounds. Agreement was reached for 16 essential and 30 recommended parameters including maternal characteristics, prenatal investigations, management and pregnancy/neonatal outcomes. Essential parameters included hypertensive complications in the current pregnancy, smoking, parity, maternal age, abdominal circumference, estimated fetal weight, umbilical artery Doppler
(pulsatility index and end-diastolic flow), middle cerebral artery Doppler, indications for intervention, pregnancy outcome (live birth, stillbirth or neonatal death), gestational age at delivery, birthweight, birthweight centile, mode of delivery and Apgar score at 5 minutes.

**Conclusions:** We present a list of essential and recommended parameters that characterize FGR independent of study hypotheses. Uniform reporting of these variables in prospective clinical research is expected to improve data quality, study consistency and ultimately our understanding of FGR.
INTRODUCTION

Fetal growth restriction (FGR) remains a major cause of adverse perinatal outcome, in particular stillbirth, neonatal death, hypoxic ischemic encephalopathy and cerebral palsy.\textsuperscript{1,2,3} FGR may arise as a consequence of several underlying causes. Even cases that are primarily due to placental dysfunction have significant variability in their phenotype. The underlying phenotype, the gestational age at which FGR is detected, and the management approach taken can all significantly impact outcome.\textsuperscript{4}

Lack of consistency in outcome reporting can be disruptive to progress in any speciality. With this in mind 78 editors of journals of Women's Health came together to form a consortium supporting the development, dissemination, and implementation of core outcome sets.\textsuperscript{5} This has triggered the development of a number of core outcome set initiatives using the Delphi method.\textsuperscript{6} These are agreed, clearly defined minimum sets of outcomes that should be measured in a standardised manner and reported consistently.\textsuperscript{7} Their development and implementation is valuable for the design and reporting of clinical studies, and ultimately aims to improve clinical care.

In addition to core outcome sets, we argue that exposure variables should also be standardized - irrespective of the primary study hypothesis. In other words, prospective clinical studies that do not consider key variables associated with FGR may be at risk of confounding and bias, as they lack contextual information for assessing the generalizability of findings. Development of a set
of core variables that should be reported in any prospective clinical research study on FGR enhances the consistency of study design independent of the primary hypothesis. This in turn would improve study quality and impact by consistently capturing essential end-points, decreasing risk of confounding and facilitating the possibility of conducting individual patient data meta-analysis and merging a number of uniformly structured datasets.

We recently used the Delphi procedure to reach consensus on the definition of the phenotype of FGR. Accordingly, it was the aim of this study to reach consensus on a list of clinical variables that should be considered essential to report for any study of pregnancies complicated by FGR. We have purposively focused on exposures and immediate/short-term outcomes.
METHODS

The methodology of the Delphi procedure as applied in this study has been previously described. Ethics approval was not required. We identified the panel members based on their publication record as lead or senior authors in studies of FGR. When inviting panel members we specifically sought wide geographic representation in order to facilitate generalizability of a minimum reporting set (MRS) of variables. We performed a search in PubMed using keyword ‘fetal growth restriction’. Authors with more than one recent (last ten years) publication were approached. Moreover, we asked the panel in the invitational email to pass the invitation to all potential experts that they thought should to be added to the list of experts. Sample sizes for Delphi studies are variable. We aimed for a sample size between 30 and 100 because this would be small enough to include only true experts and maintain speed in the process, and large enough to ensure representative pooling of judgment. By inviting this pool of experts, we were certain that we would reach this goal, anticipating drop-out. Because we found relatively few experts from Asian and African countries, we actively asked the author group in the first Delphi to recommend experts from those continents.
Data were collected in four consecutive rounds by online questionnaires that were presented to panelists through a unique token-secured link for each round. Responses were captured in Limesurvey version 2.50. Non-responders received reminder emails after two and four weeks and were excluded from subsequent survey rounds if no response was obtained. Each round included the option of offering additional items or suggestions as well as withdrawal of items from the procedure. Newly suggested items were categorized and carefully considered (i.e. discussed by the members of the Steering Group, who decided whether it would be appropriate to include them in the next round) by the panel of experts who agreed to participate for their applicability in this procedure. All additional suggested variables were included in the next round for consideration of the entire panel.

First round
A primary list of items to be considered in the MRS was presented in the following categories: maternal characteristics, prenatal investigations, clinical management and delivery/neonatal outcome. The list of variables for the first round was generated using a combination of narrative reviews and the expertise of the authors. Panel members were asked to rate each item on a 5-point Likert scale (1: very unimportant; 2: unimportant; 3: neutral; 4: important; 5: very important). They were also asked to suggest additional parameters.

Second and third rounds
In the second round, accepted and newly recommended items in round one were presented to the panel with the answer options 'yes' or 'no'. Items that in Round 1 had scored the predefined cut-off of a median Likert score of five were considered as inclusions and presented to the panel for verification for inclusion in the MRS, while items with a median score of four were presented to verify exclusion. Items with a median score of three or lower were considered rejected and verification of rejection was requested. A predefined cut-off level of 70% agreement was used to define consensus for these questions. This meant that 70% of respondents selected "yes" for inclusion of a particular variable. In the third round, parameters that fell within a 60-70% agreement were presented to the panel for re-consideration.

Final round
The purpose of this round was to categorize all items that were retained after three rounds as either ‘essential’ or ‘recommended’. ‘Essential’ parameters were those that the panel advised should be reported in all FGR studies, while ‘recommended’ parameters were those defined as important, though not mandatory for reporting. In this round, a cut-off level of 70% agreement was chosen to define consensus for these questions.

Statistical analysis

In the first round a median of Likert 5 was sufficient for inclusion, these were brought back for confirmation in the second round. Items that scored a median of Likert 4 were thought to be doubtful for inclusion and presented in the second round to vote on whether or not to include
them, with a 70% threshold for agreement. Items that scored a Likert 3 or lower were presented again to confirm rejection, also with a 70% threshold for agreement. Items that scored 60-70% agreement for inclusion in the second round were brought back for confirmation for rejection. The agreement is only for those who completed that particular round. Therefore, the exact cut off is dependent on how many panel members completed that specific round and not those at the start or endpoint.

RESULTS
87 experts, predominantly European and North American maternal-fetal medicine specialists, agreed to participate in round one of the survey (Table 1). The flow chart of the study is displayed in Figure 1. Of all invited experts, 62 panelists (71%) completed all four rounds (Supplementary material). A list of the parameters presented to the panel in the first round is shown in Supplementary Table 2. In the first round, 54 parameters scored a Likert scale of five, while 31 and 6 parameters scored a Likert of four and three, respectively (Supplementary Figures 1-4). Thirteen additional parameters were suggested (Supplementary Table 3).

In the second and third rounds, 46 items were selected for consideration as essential or recommended in the final round. With the exception of ‘maternal hypertensive disorders in pregnancy’ and ‘need for emergency cesarean section’, all items were in the original list. With the predefined threshold of agreement set at 70%, 16 of these were selected as essential in the final round (Table 2). For these variables the median agreement was 79% (range 70-100%), with a median deviation from perfect agreement of 21% (0-30%). Thirty parameters were identified as recommended (Table 3). For these, the median agreement was 53% (range 30-69%), with a median deviation from 69% agreement of 16% (0-39%).
DISCUSSION

Summary of the findings

We presented an international expert panel with a range of variables that may be captured in clinical studies and utilized the Delphi procedure to reach consensus on parameters that are considered essential; and those that are non-essential but recommended. More than 70% of experts agreed that these variables were essential. A more inclusive set of recommended parameters expands significantly on maternal characteristics, diagnostic and surveillance parameters, as well as delivery circumstances. More than 60% of experts agreed on these variables. These findings suggest that recommended variables may be of greater value for selected research questions.

Strengths and Limitations

The Delphi methodology is widely used in the development of core outcome sets and when trying to reach a consensus. However, there are a number of limitations. Firstly, we did not have data on the demographic characteristics of the experts who chose not to participate and we did not investigate the reasons for dropping out. One probable reason is that healthcare professionals might be overwhelmed by the large number of surveys they receive; several were issued at the same time. However, a response rate of 71% (62 out of 87) is considered acceptable for a Delphi consensus. Notably, the distribution of participants and level of expertise was similar in the final round in comparison to the first round. This means that the drop-out did not influence the global coverage, nor the level of experience of the panel. We had the highest
level of drop-out during the holiday seasons. When contacting the participants by phone, most participants informed us that they simply forgot to complete the survey. Moreover, in the free text answers, we did not get any signal that the experts who dropped out were critical of the process or unhappy about the attempt to gain consensus. Secondly, despite our efforts to have an international expert representation, 76% were from Europe and North America, which may be explained by the volume of publications on fetal growth restriction originating from these two continents. Moreover, the overwhelming majority of the experts who participated were European. Given the differences in practice in Europe and the United States, in particular regarding the use of Doppler parameters, the potential implications on clinical practice, research studies and training of obstetricians and sonographers should be considered.

**Interpretation of findings and comparison with existing literature**

The present study provides a list of variables that should represent the minimal reporting set for prenatally initiated research studies of FGR.\(^9\)-\(^{17}\) Regarding the items related to the prenatal investigations that were considered essential to report, the role of fetal and uterine artery Doppler has recently gained attention, both as markers of failure to reach growth potential and as prognostic markers for both short- and long-term perinatal outcomes.\(^{18}\)-\(^{22}\) In a recent meta-analysis abnormal cerebroplacental ratio (CPR) in small fetuses was associated with an increased risk of cesarean section for fetal distress (OR 7.4), low 5-min Apgar score (OR 6.9), neonatal unit admission (OR 13.0) and neonatal complications (OR 20.4).\(^{22}\) One potential explanation of the fact that the expert panel did not vote for the uterine artery Doppler to be essential, as its main role is in the identification of the pregnancies at risk of fetal growth restriction secondary to placental insufficiency, rather than influencing the management or the
timing of delivery of these pregnancies. Another interesting observation is that the ductus venosus Doppler was not prioritized by the expert panel. The most likely explanation is that the significance of the ductus venosus Doppler is only relevant to early-onset FGR. Another explanation is that the practice of monitoring these complicated pregnancies differs, with some fetal medicine experts rely on fetal Doppler, while others use CTG (or computerized CTG), and others favour the use of biophysical profile.

**Clinical and Research Implications**

Heterogeneity of study design and outcome reporting is frequent in FGR research, limiting their generalizability.\textsuperscript{23-26} We present a set of variables to be utilized by researchers in their study design. The range of variables to be selected from the recommended list will depend on the desired precision of ascertainment. The inclusion of the essential variables is likely to improve the study data quality of future studies in this field and allow robust assessment for heterogeneity when assessed in the context of meta-analyses. Having standardized parameters also allows more in-depth meta-analysis by allowing meta-regression procedures on potential covariates with relevant effects on management or diagnostic performance. Any variables required as part of a specific study question must of course be included in addition to these essential and recommended variables.

By using expert opinions in a semi-anonymous (the panel members were aware of the list of participants, as it was provided to ask for additional experts but the individual answers were anonymous) Delphi method each vote carries equal weight so can provide specific information while suppressing dominant individual opinions. In the Delphi method, repeated rounds of voting
on items by the expert panel are organized and structurally fed back in increasing detail until a consensus is achieved. The Delphi process encourages experts to reconsider their opinion in response to the group answers in every subsequent round. Delphi procedures are widely used to reach consensus definitions or core outcome sets.27,28

In this study we have purposely focused on exposures and immediate/short-term outcomes. A forthcoming and complementary Delphi procedure is being performed to reach consensus among a larger group of stakeholders (including neonatologists, midwives, general obstetricians, fetal medicine experts and patient representatives) around a broader set of outcome parameters, such as neonatal morbidity and child health and development, that need to be reported in a core outcome set (COS). The outcome parameters discussed and voted upon in the MRS procedure are summarized here and have been used to inform the first round of the complementary COS procedure, alongside the literature review performed for the COS.

A strategy for the standardization of preeclampsia research study design was published 3 years ago,29 but a similar strategy focusing on FGR is yet to be established. Adherence to standardized protocols is likely to hasten our understanding of the etiology of FGR and development of effective treatment strategies. This set of variables, when combined with the Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT) statement,30,31 should enable the researchers to design high quality FGR research studies. The team of involved key stakeholders has the potential to set up an international network, which could be a potent vehicle for the development of international guidelines, registries and setting research priorities for FGR.
Conclusion

This study provides a list of variables that are useful to describe key features of FGR. The uniform ascertainment of these variables, independent of the specific hypothesis to be tested, can potentially significantly decrease heterogeneity of prospective research. Accordingly, we recommend utilization of these variables in future publications on FGR.

Acknowledgments: We would like to acknowledge the contribution of Frans Erich, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands for his help with the Limesurvey. A list of the experts who participated in the Delphi procedure is included in Supplementary Material.

DISCLOSURE STATEMENT:
The authors report no conflict of interest
REFERENCES

5. Khan KS. The CROWN initiative: journal editors invite researchers to develop core outcomes in women’s health. *BJOG* 2014; **121**: 1181-2.


Table 1. Characteristics of the expert panel members who completed Round 1 (n=87).

<table>
<thead>
<tr>
<th>Characteristics of participants</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continent of practice</td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>56</td>
</tr>
<tr>
<td>North America</td>
<td>20</td>
</tr>
<tr>
<td>South America</td>
<td>8</td>
</tr>
<tr>
<td>Africa</td>
<td>1</td>
</tr>
<tr>
<td>Asia</td>
<td>3</td>
</tr>
<tr>
<td>Oceania</td>
<td>11</td>
</tr>
<tr>
<td>Level of experience / academic rank</td>
<td></td>
</tr>
<tr>
<td>Professor</td>
<td>48</td>
</tr>
<tr>
<td>Associate Professor / Assistant Professor</td>
<td>24</td>
</tr>
<tr>
<td>Consultant/ Specialist</td>
<td>22</td>
</tr>
<tr>
<td>Registrar/Trainee</td>
<td>1</td>
</tr>
<tr>
<td>Non-clinical investigator/researcher</td>
<td>3</td>
</tr>
<tr>
<td>Speciality</td>
<td></td>
</tr>
<tr>
<td>General gynaecologist</td>
<td>2</td>
</tr>
<tr>
<td>Maternal-Fetal Medicine Specialist</td>
<td>87</td>
</tr>
<tr>
<td>Subspecialty Fellow</td>
<td>3</td>
</tr>
<tr>
<td>Neonatologist</td>
<td>1</td>
</tr>
<tr>
<td><strong>Pathologist / Fetal Physiologist / Radiologist / Epidemiologist</strong></td>
<td>6</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>---</td>
</tr>
<tr>
<td><strong>Years of practice in current function</strong></td>
<td></td>
</tr>
<tr>
<td>0-4</td>
<td>11</td>
</tr>
<tr>
<td>5-9</td>
<td>20</td>
</tr>
<tr>
<td>10-20</td>
<td>38</td>
</tr>
<tr>
<td>&gt;20</td>
<td>31</td>
</tr>
<tr>
<td><strong>Level of care</strong></td>
<td></td>
</tr>
<tr>
<td>Non-fetal maternal center</td>
<td>5</td>
</tr>
<tr>
<td>Fetal medicine center offering prenatal diagnosis but no fetal therapy</td>
<td>26</td>
</tr>
<tr>
<td>Fetal medicine center offering prenatal diagnosis and fetal therapy</td>
<td>63</td>
</tr>
<tr>
<td>Private hospital/tertiary care facility</td>
<td>6</td>
</tr>
<tr>
<td><strong>Referral centre for pregnancies complicated by fetal growth restriction</strong></td>
<td>92</td>
</tr>
</tbody>
</table>

**Deliveries**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1000-2500</td>
<td>28</td>
</tr>
<tr>
<td>2500-5000</td>
<td>31</td>
</tr>
<tr>
<td>&gt;5000</td>
<td>41</td>
</tr>
</tbody>
</table>
Table 2. Parameters identified as essential by consensus.

<table>
<thead>
<tr>
<th>Minimum reporting set (MRS) essential parameters</th>
<th>% consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maternal characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>• Hypertensive complications in current pregnancy</td>
<td>92</td>
</tr>
<tr>
<td>• Smoking</td>
<td>79</td>
</tr>
<tr>
<td>• Parity</td>
<td>79</td>
</tr>
<tr>
<td>• Maternal age</td>
<td>77</td>
</tr>
<tr>
<td><strong>Items for investigations/assessment</strong></td>
<td></td>
</tr>
<tr>
<td>• Abdominal circumference</td>
<td>89</td>
</tr>
<tr>
<td>• Estimated fetal weight</td>
<td>89</td>
</tr>
<tr>
<td>• Umbilical artery Doppler index</td>
<td>82</td>
</tr>
<tr>
<td>• Umbilical artery end-diastolic flow</td>
<td>74</td>
</tr>
<tr>
<td>• Middle cerebral artery Doppler</td>
<td>70</td>
</tr>
<tr>
<td><strong>Items for management</strong></td>
<td></td>
</tr>
<tr>
<td>• Indications for intervention</td>
<td>72</td>
</tr>
<tr>
<td><strong>Pregnancy outcome</strong></td>
<td></td>
</tr>
<tr>
<td>• Gestational age at delivery</td>
<td>100</td>
</tr>
<tr>
<td>• Birthweight</td>
<td>98</td>
</tr>
<tr>
<td>• Pregnancy outcome (livebirth, stillbirth, intrapartum or neonatal death)</td>
<td>97</td>
</tr>
<tr>
<td>• Mode of delivery</td>
<td>79</td>
</tr>
<tr>
<td>• Apgar score at 5 minutes</td>
<td>74</td>
</tr>
<tr>
<td>Birthweight centile</td>
<td>72</td>
</tr>
</tbody>
</table>
Table 3. Parameters identified as recommended by consensus.

<table>
<thead>
<tr>
<th>Parameters recommended to report</th>
<th>% consensus</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Maternal characteristics</strong></td>
<td></td>
</tr>
<tr>
<td>• Maternal body mass index</td>
<td>69</td>
</tr>
<tr>
<td>• Prior fetal growth restriction or small baby</td>
<td>61</td>
</tr>
<tr>
<td>• Maternal ethnic background</td>
<td>52</td>
</tr>
<tr>
<td>• Drug/alcohol abuse</td>
<td>51</td>
</tr>
<tr>
<td>• Prior essential hypertension</td>
<td>49</td>
</tr>
<tr>
<td>• Prior preeclampsia</td>
<td>46</td>
</tr>
<tr>
<td>• Prior autoimmune disease</td>
<td>39</td>
</tr>
<tr>
<td>• Prior stillbirth</td>
<td>39</td>
</tr>
<tr>
<td>• Pre-existing diabetes</td>
<td>37</td>
</tr>
<tr>
<td>• Pre-existing vascular disease</td>
<td>36</td>
</tr>
<tr>
<td>• Pre-existing renal impairment</td>
<td>36</td>
</tr>
<tr>
<td>• Mode of conception</td>
<td>34</td>
</tr>
<tr>
<td><strong>Items for investigations/ assessment</strong></td>
<td></td>
</tr>
<tr>
<td>• Growth centile formula</td>
<td>64</td>
</tr>
<tr>
<td>• Femur length</td>
<td>61</td>
</tr>
<tr>
<td>• Head circumference</td>
<td>57</td>
</tr>
</tbody>
</table>
- Amniotic fluid index / single deepest pocket 54
- Uterine artery Doppler index 41
- Fetal karyotype 33

**Items for management**

- Interval between last examination and delivery 69
- Method of monitoring 64
- Administration of steroids to promote fetal lung maturity 64
- Intervals between assessments in longitudinal studies 57
- Administration of low dose aspirin 54
- Frequency of monitoring 44
- Administration of MgSO4 30

**Pregnancy outcome**

- Need for emergency cesarean section 66
- Fetal sex 62
- Umbilical artery pH 59
- Signs of fetal distress/hypoxia on fetal heart rate monitoring 56
- Onset of labor 38
FIGURE LEGENDS

Figure 1. The Delphi flow chart

Supplementary Figure 1. Likert scores of the parameters describing maternal characteristics

Supplementary Figure 2. Likert scores of parameters describing investigations/assessment

Supplementary Figure 3. Likert scores of parameters describing management

Supplementary Figure 4. Likert scores of parameters describing pregnancy outcome (S4a) and neonatal outcomes (S4b)
Definition of problem

Creation of the expert panel
Invitation by literature research and recommendation (N=100)

Round 1 (N=87)
- Scoring on 5-point Likert scale of proposed items
  (5: Inclusion, 4: debatable for inclusion, <= 3: rejection)
- Suggestion of additional items by the panel

Round 2 (N=79)
- Double check inclusion items Likert 4 & 5 and exclusion Likert <=3
- Introduce suggested parameters for scoring
  Cut-off level consensus 70% agreement

Round 3 (N=73)
- Double check exclusion items 60-70% agreement
  Cut-off level consensus 70% agreement

Final round (N=62)
- Differentiation between:
  o Essential items: the final MRS
  o Recommended items to report

Figure 1.tif