Author's response to reviews

Title: Facilitators and barriers to facility-based delivery in low- and middle-income countries: A qualitative evidence synthesis

Authors:

Meghan A Bohren (mbohren1@jhu.edu)
Erin C Hunter (erinchunter@gmail.com)
Heather M Munthe-Kaas (Heather.Munthe-Kaas@kunnskapssenteret.no)
João Paulo Souza (jpsouza@fmrp.usp.br)
Joshua P Vogel (vogeljo@who.int)
A. Metin Gülmezoglu (gulmezoglu@who.int)

Version: 3 Date: 18 June 2014

Author's response to reviews: see over
Reviewer's report
Title: Facilitators and barriers to facility-based delivery in low- and middle-income countries: A qualitative evidence synthesis
Version: 1 Date: 8 April 2014
Reviewer: Christopher Carroll

Reviewer's report:
I think this is an interesting review, which some well-conducted elements; however, there are some major issues with the reporting and justification of the processes undertaken.

Major compulsory revisions:
1. The rationale for the study is not explicit from the Background. It states that, to reduce maternal and perinatal mortality rates, an intervention of known effect is "skilled attendance at facility-based deliveries (6-8). Skilled providers attend 99.5% of births in more developed countries, but only 61.9% in less developed countries and 35.3% in the least developed countries (6)". There are two issues here. First, are these providers all in "facilities" or are they in the community? This is unclear. Indeed, the "catch-all" nature of the term is only mentioned later, in the Limitations. The terms "health-facilities" and "facility-based delivery" need to be clearly defined at the very start. Second, the authors present no evidence that there is a real issue with women attending facilities / accessing this intervention; it is implied, but not demonstrated. If this evidence was provided, then this would justify more clearly the case for the synthesis of evidence on people's views regarding the intervention.

Thank you for your feedback. We added text to clarify the scope of the problem and provided definitions in the introduction:

“According to UNICEF’s 2014 estimates, facility-based delivery rates remained disappointingly low in several regions, including 48% in sub-Saharan Africa, 44% in South Asia, and 71% in the Middle East and North Africa (9). Globally, in the least developed countries, facility-based delivery rates in 2014 averaged at 43% (9).”

“For the purpose of this review, we have defined a facility-based delivery as a birth occurring in health facility of any level from community health center through tertiary facility.”

“While population-based surveys capture important information regarding the proportion of births occurring in health facilities, surveys are unable to capture the complex myriad explanations for women’s health practices and preferences in delivery location. Qualitative research methods are therefore useful complements to population-based surveys to understand how women perceive, interpret, and weigh a range of factors that affect their delivery location. Synthesizing qualitative evidence allows us to aggregate explanations of the “how” and the “why” behind the decision to deliver at a facility or elsewhere across multiple contexts.”
2. Synthesis: If I have understood correctly, there was a single round of coding to identify the framework themes, and then a second round to populate that framework? Did all of the data fit into the framework produced by the initial open codes? There were no data that did not fit under existing themes/sub-themes? This seems surprising. thematic analysis tends to be more iterative than this. No citation is given for the process of thematic analysis used, other than the 1994 ATLAS software. More up-to-date methodological references are required, especially as thematic synthesis is the the common, published approach.

The initial round of coding by two reviewers developed the themes presented in Box 1. All text units were subsequently classified into one of the themes. The initial coding scheme was intended to be general and very broad, in order to capture all the overarching core themes present in the data.

Then, all text units within each broad theme were further analyzed by two reviewers to develop the hierarchical [axial] coding scheme presented in Appendix C: Codebook. Appendix C presents a description of each hierarchical axial code and an illustrative quote, as well as demonstrates the relationship between the axial codes. Axial coding is widely accepted in qualitative literature as a sufficient method to disaggregate core themes during qualitative analysis (Strauss and Corbin 1990, Strauss and Corbin 1998, Charmaz 2006).

[Please see the text for further elucidation of the analysis methods.]

3. Assessing certainty of findings: CerQual is a new approach and its mechanics really need explaining more fully here: is CASP being used to assess how well studies are conducted? How is coherence assessed? I like Table 2 - the results of the synthesis and application of CerQual; I can see exactly how/why a finding has been weighted like it has.

Thank you for your feedback. We further elucidated the CERQual approach in the text:

“First, we appraised the methodological quality of the individual studies contributing to each review finding using the modified CASP tool discussed previously. The methodological assessment of the individual studies contributing to each review finding is important to determine how likely it is that the research produced credible results, how precise and dependable an understanding of the phenomenon of interest the research will provide, and how widely the research findings could be applied. In the CERQual approach, confidence in a review finding is weakened when the primary studies that contribute to each review finding have critical methodological weaknesses. Second, we assessed the coherence of each review finding by exploring to what extent clear patterns could be identified across the data contributed by each of the individual studies, or that plausible explanations are provided if there is variation across individual studies. Assessing the coherence in each review finding is important to ensure that
each review finding is grounded in patterns found in the data. The main threat to the coherence of a review finding is unexplained inconsistencies found from variations in the data from individual studies. Based on the assessment of the methodological quality of individual studies contributing to each review finding and the coherence of each review finding, the confidence in the evidence for each review finding was assessed as high, moderate, and low (table 2)."

4. An edited and abbreviated version of Appendix F, a summary of the included studies, is really needed in the text, at the start of the Results.

We abbreviated the appendix and added it as a table in the text.

5. I am unclear how the authors progressed from the findings in Table 2, to what is presented in Figure 2. Please explain.

Thank you for your feedback – great point. We added an explanatory caption to figure 2: “Figure 2. Multi-level life course framework of facility-based delivery in LMICs. This framework was developed using the multi-level life course approach to explore how experiences earlier in an individual’s life impact their subsequent decisions and actions, and how these experiences range across individual, family, community, and national level influences. The framework was developed after the review findings and first, second, and third order themes were finalized.”

6. Limitations:
The whole social, economic and health system context can be very different in different LMICs, and this raises questions about grouping studies together from so many different countries. Can people in LMICs all be grouped together, treated as a homogenous population for a synthesis such as this? There is also no attempt to conduct any form of sensitivity analysis to gauge the impact of findings from different countries and contexts, or to group findings by types of intervention (e.g. physical facilities vs alternatives). Thirty-four is a large number of studies for a QES and some form of sub-grouping would have made sense.

Thank you for your feedback, and we agree that the social, economic and health system contexts in LMICs vary greatly and could impact health beliefs and behaviors differently. This review was commissioned by the World Health Organization as a global systematic review of facilitators and barriers to facility-based deliveries, as part of the global research agenda to improve intrapartum care. As such, we were specifically requested to explore barriers and facilitators to facility-based childbirth across all LMICs. In an attempt to address potential differences across contexts, we used several approaches. First, we sought to develop a multi-level life course conceptual framework that would be applicable across multiple settings but, as necessary, could be adapted to a specific context (i.e.: regional, national, state-level). Second, we conducted a sensitivity analysis to determine if there were regional differences in the review findings and labeled every
text unit with the region from which the data came (in this review, most studies were from Africa or Asia, with a few studies from Middle East and Central/South America). In general, we found that most of the review findings were cross-cutting; therefore, it was more logical to present the findings at a global level. An example of where regional themes were divergent was in the section on HIV stigma (review findings #36, 37, 38). The only included papers that discussed fears related to HIV status and testing at the facility as a barrier to facility-based delivery were all from Kenya. Therefore, the explanation of certainty in the evidence assessments state that while these review findings are of low certainty at the global level of analysis, it is possible that the certainty of the finding could be higher in Kenya (or in contexts similar to Kenya).

Minor essential revisions:
1. The type of search undertaken needs justifying:
The search seems completely overblown; a mammoth task, and very unusual in the current methodological climate, which tends towards sampling (Suri 2011) and other purposive or focused approaches for qualitative evidence synthesis, rather than the traditional Cochrane quantitative review, highly-sensitive search strategy approach.
It is also unclear how such a search translated to the non-standard bibliographic resources, such as WHO and Google Scholar; this is not explained (nor is it reproducible).
Why were the references of included studies not searched? Or citation searching of included studies? Which are standard, and potentially more effective and far less resource-intensive and time-consuming means of identifying relevant papers.
Also, why wasn’t a published filter used for qualitative research (e.g. Wilczynski 2007 for CINAHL)? Also, the filter that was employed includes repetition of the same term in different search strings, e.g. "focus group" in #24 and #26

Thank you for this very helpful feedback on the search strategy, and we agree that the discussion related to highly sensitive Cochrane-esque search strategies versus purposive sampling in qualitative evidence synthesis is ongoing. This review was commissioned by the World Health Organization as a global systematic review of facilitators and barriers to facility-based deliveries, as part of the global research agenda to improve intrapartum care. As such, we sought to use a comprehensive and sensitive search strategy that we developed with the assistance of the WHO librarian and the Norwegian Knowledge Center for the Health Sciences librarian. We were not comfortable limiting our search to purposive sampling as this may have restricted the studies that we found to a certain region or context that may not have been appropriate for this review. The references of the included studies were hand searched to identify additional references, and this is now clarified in the “search strategy” section of the methods. We also personally contacted several researchers in relevant fields for their assistance in identifying relevant studies:

“We also searched WHO Global Health Library, Cochrane Library, DARE, Google Scholar, CRD, OpenGrey, and EThOs for gray literature and unpublished reports. We
also personally contacted researchers in relevant fields of study for assistance in identifying studies. The reference lists of all included studies were hand searched to identify any potentially relevant studies.”

Thank you also for providing the citation to the CINAHL qualitative search filter (Wilczynski 2007), which we will most definitely consider in future qualitative evidence syntheses.

2. Are ENTREQ rather than - or on addition to PRISMA - not more appropriate guidelines for reporting a QES?

http://www.biomedcentral.com/1471-2288/12/181

This is a very helpful suggestion since, at the time we conducted our review, none of the authors were familiar with ENTREQ. Please find the ENTREQ statement below with the corresponding page numbers where our manuscript addresses these issues:

<table>
<thead>
<tr>
<th>#</th>
<th>Item</th>
<th>Guide and description</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aim</td>
<td>State the research question the synthesis addresses.</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Synthesis methodology</td>
<td>Identify the synthesis methodology or theoretical framework which underpins the synthesis, and describe the rationale for choice of methodology (e.g. meta-ethnography, thematic synthesis, critical interpretive synthesis, grounded theory synthesis, realist synthesis, meta-aggregation, meta-study, framework synthesis).</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Approach to searching</td>
<td>Indicate whether the search was pre-planned (comprehensive search strategies to seek all available studies) or iterative (to seek all available concepts until theoretical saturation is achieved).</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Inclusion criteria</td>
<td>Specify the inclusion/exclusion criteria (e.g. in terms of population, language, year limits, type of publication, study type).</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Data sources</td>
<td>Describe the information sources used (e.g. electronic databases (MEDLINE, EMBASE, CINAHL, psycINFO, Econlit), grey literature databases (digital thesis, policy reports), relevant organisational websites, experts, information specialists, generic web searches (Google Scholar) hand searching, reference lists) and when the searches conducted; provide the rationale for using the data sources.</td>
<td>3-4</td>
</tr>
<tr>
<td>6</td>
<td>Electronic Search strategy</td>
<td>Describe the literature search (e.g. provide electronic search strategies with population terms, clinical or health topic terms, experiential or social phenomena related terms, filters for qualitative research, and search limits).</td>
<td>30-37</td>
</tr>
<tr>
<td>7</td>
<td>Study screening methods</td>
<td>Describe the process of study screening and sifting (e.g. title, abstract and full text review, number of independent reviewers who screened studies).</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Study characteristics</td>
<td>Present the characteristics of the included studies (e.g. year of publication, country, population, number of participants, data collection, methodology, analysis, research questions).</td>
<td>Appendix F, 43-49</td>
</tr>
<tr>
<td>9</td>
<td>Study selection results</td>
<td>Identify the number of studies screened and provide reasons for study exclusion (e.g. for comprehensive searching, provide numbers of studies screened and reasons for exclusion indicated in a figure/flowchart; for iterative searching describe reasons for study exclusion and inclusion based on modifications to the research question and/or contribution to theory development).</td>
<td>Figure 1</td>
</tr>
<tr>
<td>10</td>
<td>Rationale for appraisal</td>
<td>Describe the rationale and approach used to appraise the included studies or selected findings (e.g. assessment of conduct (validity and robustness), assessment of reporting (transparency), assessment of content and utility of the findings).</td>
<td>4-6</td>
</tr>
</tbody>
</table>
3. The Discussion does not appear to consider specifically the "weighting" of findings, as expressed in Table 2. It would help to have a discussion of which findings have high certainty, and which moderate or low, to inform the conclusions.

Thank you for your feedback. Our primary aim in the discussion was to elucidate the newer or under-explored areas of research and practice that this review brought to light. While we recognize the importance of accessibility issues (i.e.: travel to facilities and prohibitive costs), these issues have been highlighted by research and practice in the past few decades since Thaddeus and Maine’s three delays model. We believe that one of the strengths of this review will be to help bring issues such as “improving intrapartum care” more firmly into the academic discourse on the seeking of facility-based delivery, and thus draw more attention to addressing these existing barriers to utilization of maternal health resources (and specifically to facility-based delivery). To this end, we chose to use the allot word limit to emphasize the dearth of research related to mistreatment of women and perceived low quality of care from the user perspective during the intrapartum period, with the hope that highlighting this important issue would generate more momentum around the topic. We added text to the limitations section to address where data might have been sparse.

**Level of interest:** An article whose findings are important to those with closely
related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
I declare that I have no competing interests
Reviewer's report
Title: Facilitators and barriers to facility-based delivery in low- and middle-income countries: A qualitative evidence synthesis
Version: 1 Date: 1 May 2014
Reviewer: Karin Hannes

Reviewer's report:
Review report Bohren et al.
Overall comment:
This is a well written report that follows the current state of the art in the methodology of systematic reviews. However, the information on the methodology used is rather scarce. The procedure could better be explained so that starting reviewers can draw from this part for their own projects.

- Major Compulsory Revisions
  • Methods section: Quality Assessment: a quality assessment has been done, but it is not fully clear whether studies were excluded based on lower quality or what has been done with potentially flawed paper findings in the findings section. This should be explained or shown through, for example, a sensitivity analysis.

Thank you for your feedback. No studies were excluded as a result of the quality assessment; rather, this provided an assessment of weaknesses in study methodologies which enabled us to better evaluate the review findings. In particular, the quality assessment played a critical role in assessing the certainty of the findings. We added text to the “Assessing the certainty of the findings” section to better explain what role the quality assessments played in the synthesis:

“First, we appraised the methodological quality of the individual studies contributing to each review finding using the modified CASP tool discussed previously. The methodological assessment of the individual studies contributing to each review finding is important to determine how likely it is that the research produced credible results, how precise and dependable an understanding of the phenomenon of interest the research will provide, and how widely the research findings could be applied. In the CERQual approach, confidence in a review finding is weakened when the primary studies that contribute to each review finding have critical methodological weaknesses. Second, we assessed the coherence of each review finding by exploring to what extent clear patterns could be identified across the data contributed by each of the individual studies, or that plausible explanations are provided if there is variation across individual studies. Assessing the coherence in each review finding is important to ensure that each review finding is grounded in patterns found in the data. The main threat to the coherence of a review finding is unexplained inconsistencies found from variations in the data from individual studies. Based on the assessment of the methodological quality of individual studies contributing to each review finding and the coherence of each review finding, the confidence in the evidence for each review finding was assessed as high, moderate, and low (table 2).”

• The outline of the synthesis method is fairly limited and does not provide much detail on what has been going on in practice. This section should be expanded
on. It does not help other authors to repeat the study. It is like the themes and subcategories suddenly ‘emerge’ from the data; make this move from text units to categories and themes more explicit. I can imagine that the use of 1st, 2nd and 3th order concepts, for example, may help.

Thank you for your suggestion. We added text to the “Synthesis” section of the methods in order to provide additional detail on how the coding structure/themes were developed.

The initial round of coding developed the themes presented in Box 1. All text units were subsequently classified into one of the themes. The initial coding scheme was intentionally general and very broad in order to capture all the overarching core themes present in the data. Then, all text units within each theme were further analyzed to develop the axial coding scheme (appendix C). Axial coding is widely accepted in qualitative literature as a sufficient method to disaggregate core themes during qualitative analysis (Strauss and Corbin 1990, Strauss and Corbin 1998, Charmaz 2006). MB and EH applied the axial codes systematically to the data by hand-sorting all the individual text units into themes and sub-themes.

• Certainty of findings: explain how this differs from the CASP, which is also a critical appraisal exercise including aspects of evaluating how certain data are or how much trust we can put into them.

Thank you – We have added text to the “assessing the certainty of the findings” section of the methods to address this point:

“First, we appraised the methodological quality of the individual studies contributing to each review finding using the modified CASP tool discussed previously. The methodological assessment of the individual studies contributing to each review finding is important to determine how likely it is that the research produced credible results, how precise and dependable an understanding of the phenomenon of interest the research will provide, and how widely the research findings could be applied. In the CERQual approach, confidence in a review finding is weakened when the primary studies that contribute to each review finding have critical methodological weaknesses. Second, we assessed the coherence of each review finding by exploring to what extent clear patterns could be identified across the data contributed by each of the individual studies, or that plausible explanations are provided if there is variation across individual studies. Assessing the coherence in each review finding is important to ensure that each review finding is grounded in patterns found in the data. The main threat to the coherence of a review finding is unexplained inconsistencies found from variations in the data from individual studies. Based on the assessment of the methodological quality of individual studies contributing to each review finding and the coherence of each review finding, the confidence in the evidence for each review finding was assessed as high, moderate, and low (table 2).”
The PRISMA statement has been developed for quantitative reviews. Did you add any specific criteria to it that better matches what should be reported in QES. The ENTREQ statement has been developed to this end. You may want to match your report with these criteria.

This is a very helpful suggestion since, at the time we conducted our review, none of the authors were familiar with ENTREQ. Please find the ENTREQ statement below with the corresponding page numbers where our manuscript addresses these issues:

<table>
<thead>
<tr>
<th>#</th>
<th>Item</th>
<th>Guide and description</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Aim</td>
<td>State the research question the synthesis addresses.</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>Synthesis methodology</td>
<td>Identify the synthesis methodology or theoretical framework which underpins the synthesis, and describe the rationale for choice of methodology (e.g. meta-ethnography, thematic synthesis, critical interpretive synthesis, grounded theory synthesis, realist synthesis, meta-aggregation, meta-study, framework synthesis).</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>Approach to searching</td>
<td>Indicate whether the search was pre-planned (comprehensive search strategies to seek all available studies) or iterative (to seek all available concepts until they theoretical saturation is achieved).</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Inclusion criteria</td>
<td>Specify the inclusion/exclusion criteria (e.g. in terms of population, language, year limits, type of publication, study type).</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Data sources</td>
<td>Describe the information sources used (e.g. electronic databases (MEDLINE, EMBASE, CINAHL, psycINFO, Econlit), grey literature databases (digital thesis, policy reports), relevant organisational websites, experts, information specialists, generic web searches (Google Scholar) hand searching, reference lists) and when the searches conducted; provide the rationale for using the data sources.</td>
<td>3-4</td>
</tr>
<tr>
<td>6</td>
<td>Electronic Search strategy</td>
<td>Describe the literature search (e.g. provide electronic search strategies with population terms, clinical or health topic terms, experiential or social phenomena related terms, filters for qualitative research, and search limits).</td>
<td>30-37</td>
</tr>
<tr>
<td>7</td>
<td>Study screening methods</td>
<td>Describe the process of study screening and sifting (e.g. title, abstract and full text review, number of independent reviewers who screened studies).</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Study characteristics</td>
<td>Present the characteristics of the included studies (e.g. year of publication, country, population, number of participants, data collection, methodology, analysis, research questions).</td>
<td>Appendix F, 43-49</td>
</tr>
<tr>
<td>9</td>
<td>Study selection results</td>
<td>Identify the number of studies screened and provide reasons for study exclusion (e.g. for comprehensive searching, provide numbers of studies screened and reasons for exclusion indicated in a figure/flowchart; for iterative searching describe reasons for study exclusion and inclusion based on modifications to the research question and/or contribution to theory development).</td>
<td>Figure 1</td>
</tr>
<tr>
<td>10</td>
<td>Rationale for appraisal</td>
<td>Describe the rationale and approach used to appraise the included studies or selected findings (e.g. assessment of conduct (validity and robustness), assessment of reporting (transparency), assessment of content and utility of the findings).</td>
<td>4-6</td>
</tr>
<tr>
<td>11</td>
<td>Appraisal items</td>
<td>State the tools, frameworks and criteria used to appraise the studies or selected findings (e.g. Existing tools: CASP, QARI, COREQ, Mays and Pope [25]; reviewer developed tools; describe the domains assessed: research team, study design, data analysis and interpretations, reporting).</td>
<td>4-6</td>
</tr>
<tr>
<td>12</td>
<td>Appraisal process</td>
<td>Indicate whether the appraisal was conducted independently by more than one reviewer and if consensus was required.</td>
<td>4-6</td>
</tr>
<tr>
<td>13</td>
<td>Appraisal results</td>
<td>Present results of the quality assessment and indicate which articles, if any, were weighted/excluded based on the assessment and give the rationale.</td>
<td>4-6, Table 2</td>
</tr>
<tr>
<td>14</td>
<td>Data extraction</td>
<td>Indicate which sections of the primary studies were analysed and how were the data extracted from the primary studies? (e.g. all text under the headings “results /conclusions” were extracted electronically and entered into a computer software).</td>
<td>4</td>
</tr>
<tr>
<td>15</td>
<td>Software</td>
<td>State the computer software used, if any.</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>16</td>
<td>Number of reviewers</td>
<td>Identify who was involved in coding and analysis</td>
<td>4-6</td>
</tr>
<tr>
<td>17</td>
<td>Coding</td>
<td>Describe the process for coding of data (e.g. line by line coding to search for concepts).</td>
<td>4-6</td>
</tr>
<tr>
<td>18</td>
<td>Study comparison</td>
<td>Describe how were comparisons made within and across studies (e.g. subsequent studies were coded into pre-existing concepts, and new concepts were created when deemed necessary).</td>
<td>4-6</td>
</tr>
<tr>
<td>19</td>
<td>Derivation of themes</td>
<td>Explain whether the process of deriving the themes or constructs was inductive or deductive</td>
<td>4-6</td>
</tr>
<tr>
<td>20</td>
<td>Quotations</td>
<td>Provide quotations from the primary studies to illustrate themes/constructs, and identify whether the quotations were participant quotations of the author’s interpretation.</td>
<td>6-12, 38-42</td>
</tr>
<tr>
<td>21</td>
<td>Synthesis output</td>
<td>Present rich, compelling and useful results that go beyond a summary of the primary studies (e.g. new interpretation, models of evidence, conceptual models, analytical framework, development of a new theory or construct).</td>
<td>6-12, 25-29, figure 2, table 1, table 2</td>
</tr>
</tbody>
</table>

- Explain how box 1 has been developed. Was it an a-priori scheme based on theory or a scheme that has been generated bottom-up? What is the relation between box 1 and table 1? How has it been used in the analysis? This is unclear.

Thank you for your feedback. To address this concern, we added text to the “Synthesis” section of the methods in order to clarify our coding process:

“The initial round of coding developed the themes presented in box 1. All text units were subsequently classified into one of the themes. The initial coding scheme was intentionally general and broad in order to capture the overarching core themes present in the data. Then, each theme was further analyzed to develop the axial coding scheme (appendix C). Axial coding is widely accepted in qualitative literature as a sufficient method to disaggregate core themes during qualitative analysis (Strauss and Corbin 1990, Strauss and Corbin 1998, Charmaz 2006). MB and EH applied the axial codes systematically to the data by hand-sorting the text units into themes and sub-themes. Table 1 pictographically presents the first, second, and third order themes that emerged from the initial and axial coding. First order themes represent text units that are grouped together based on common themes. Second order themes represent first order themes grouped together based on common, higher-level themes. Third order themes represent overarching high-level themes comprised of the first- and second-level themes (Braun & Clarke 2006).”

Minor Essential Revisions

- Background section:
  Potential correction: I don’t think that the use of systematic methods reduces bias persé. I think it increases transparency, which allows the reader to better judge the credibility and trustworthiness of Qualitative Evidence Syntheses (QES).
We agree with this suggestion and have revised the text to say: “Approaching qualitative evidence synthesis using systematic methodologies increases the transparency, credibility, trustworthiness, and confidence in each of the review findings.

• Methods section:
  Data-extraction: How many authors responded on the request for information? What form or which descriptive data were used for the extraction of data?

  All authors who were contacted (4 in total) responded to requests for further information. Requests for further information generally were in relation to qualitative analysis method employed by the primary authors.

  A data extraction form developed by MB, HMK, and EH was used in this review. The data extraction form contained the following main domains:
  • Study setting
  • Demographics
  • Study objectives
  • Study design
  • Data collection methods
  • Data analysis methods
  • Themes and text related to the review question
  • Study conclusions

  • Findings: Change review into findings. You methods part is also part of the review. Number the studies in appendix E or mark the included studies in the reference list with an *.

  We revised the title to “Findings.”

  We added numbers to the studies listed in Appendix F “Study summaries”. All studies (34) summarized in this appendix are included in the review.

  • Provide study numbers for the statement about elder woman pressuring younger women to deliver at home. In the effects to policy section, link study numbers to subdivisions in the phrase.

  We added study numbers to elder women and policy sections.

  • Explain the labels in table 2: what is certainty in the evidence? The 1st and the 3th account in the table have a different certainty label, however have the same explanation assigned to it. The 11th and 12th account, for example have the
same certainty label, however a different explanation of it. This is confusing.

Thank you for your feedback. The method section now further elucidates the CERQual methodology, which is an assessment of the certainty in the evidence (please see the text in the “assessing the certainty of the findings” section of the methods). We hope this helps clarify the confusion.

As now detailed in the methods section, the CERQual methodology employs an assessment of the methodological quality of the individual studies contributing to each review finding, as well as the coherence of each review finding. For example, review finding #1 was developed from 5 primary studies, 3 of which had “high” quality assessments and 2 of which has “medium” quality assessments. This led to an overall assessment of the certainty of the evidence in this particular review finding as “moderate”. Had there been more sufficient evidence regarding this review finding from other included primary studies (for example, if 5 more “moderate” or “high” quality studies contributed to the review finding), then it is possible that this review finding could have been rated “high”.

Review finding #3 was developed from 16 primary studies, 6 of which had “high” quality assessments, 6 of which had “medium” quality assessments, and 4 of which had “low” quality assessments.

- Discretionary Revisions
It is unclear whether the appendixes will be published with the review, but it seems necessary. You may want to consider a link to an online available annex in the article.

Yes, the authors agree that the appendices should be included in the publication. There were a lot of explanatory tables, figures, and text that we included in the appendices so as to reduce clutter in the actual manuscript. We would like to work with the editors to ensure that readers have access to all available appendices, as we believe that it will help other researchers to replicate our process.

Level of interest: An article whose findings are important to those with closely related research interests
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: I declare that I have no competing interests.
Reviewer's report

Title: Facilitators and barriers to facility-based delivery in low- and middle-income countries: A qualitative evidence synthesis

Version: 1  Date: 11 May 2014

Reviewer: Shehla Zaidi

Reviewer's report:

This can make an extremely interesting paper, of good practical import for researchers and policymakers working in the area of maternal care, provided suggested further improvements are made to the paper.

1. Write-up of the Methodology needs further elaboration. The appended tables show that sufficient level of work has gone in but the narrative does not do it justice. Particularly needed are details on what basis the studies were included or excluded. The CASP quality assessment tool is mentioned but details are needed on what this is and how you adapted it.

Thank you for your interest in the review and your valuable feedback. We added text to the “quality assessment” section of the methods to further elucidate our use of the CASP checklist:

“The CASP checklist was adapted from a checklist form to a spreadsheet form that allowed for a more in-depth discussion of potential methodological challenges in the primary studies. The modified forms included the following domains: research aims, methodology, research design, recruitment strategy, data collection, data analysis, reflexivity, ethical considerations, findings, and value of research. The overall quality assessment of “high”, “medium”, or “low” was based on the evaluation by two reviewers and active discussion until consensus was reached in the case of rating discrepancies.”

Additional text has also been added to the assessment of the certainty of the findings that is relevant to the quality assessments:

“First, we appraised the methodological quality of the individual studies contributing to each review finding using the modified CASP tool discussed previously. The methodological assessment of the individual studies contributing to each review finding is important to determine how likely it is that the research produced valid results, how precise and reliable an understanding of the phenomenon of interest the research will provide, and how widely the research findings could be applied. In the CERQual approach, confidence in a review finding is weakened when the primary studies that contribute to each review finding have critical methodological weaknesses. Second, we assessed the coherence of each review finding by exploring to what extent clear patterns could be identified across the data contributed by each of the individual studies, or that plausible explanations are provided if there is variation across individual studies. Assessing the coherence in each review finding is important to ensure that each review finding is grounded in patterns found in the data. The main threat to the coherence of a review finding is unexplained inconsistencies found from variations in the data from individual studies. Based on the assessment of the methodological quality of individual studies contributing to each review finding and the coherence of each
review finding, the confidence in the evidence for each review finding was assessed as high, moderate, and low (table 2)."

Inclusion and exclusion criteria are listed in the “study selection” section of the methods and presented in Figure 1 (“The original PubMed and CINAHL search yielded 2,275 articles, from which 101 duplicates were excluded. Three reviewers (MB, EH, HMK) independently screened titles and abstracts for inclusion, then reviewed the full text articles using standardized inclusion criteria: (a) primary data; (b) English or French language; (c) LMIC; (d) study objectives related to barriers and/or facilitators to facility-based delivery; (e) qualitative data collection method; (f) qualitative analysis method; and (g) full text available. Studies that did not report qualitative data in their findings sections were excluded.”)

2. The search can also be widened with use of other search engines especially for grey literature. Just 2 search engines for a systematic review appears narrow.

Systematic searches were developed in collaboration with two librarians and were conducted in PubMed (Appendix A for full search strategy) and CINAHL (Appendix B for full search strategy). The WHO Global Health Library, Cochrane Library, DARE, Google Scholar, CRD, OpenGrey, and EThOs for gray literature and unpublished reports. The reviewers also contacted qualitative researchers in relevant fields and hand searched through the reference lists.

We added this clarification to the “search strategy” section of the methods:

“We also personally contacted researchers in relevant fields of study for assistance in identifying studies. The reference lists of all included studies were hand searched to identify any potentially relevant studies.”

3. Qualitative research needs equal rigor as quantitative research. Please show the MESH words in a table. Also what was the criteria for shortlisting, the difficulties faced in the process and how these were resolved.

Thank you for your feedback. We have made the MESH words available in the search strategies for PubMed (appendix A) and CINAHL (appendix B).

Criteria for shortlisting the included studies are outlined in the “study selection” section of the methods [“(a) primary data; (b) English or French language; (c) LMIC; (d) study objectives related to barriers and/or facilitators to facility-based delivery; (e) qualitative data collection method; (f) qualitative analysis method; and (g) full text available.”]. The main difficulty we faced in shortlisting studies for inclusion was that a couple of potentially relevant primary studies identified through the various search strategies did not include any description of their analysis methods. Given that part of our inclusion
criteria was “use of qualitative data collection and qualitative data analysis methods”, we were concerned that the requirement of a description of qualitative analysis methods could be too strict. However, we ultimately decided that the inclusion criteria for qualitative analysis method was important, as several studies used quantitative analysis methods to analyze qualitative data (i.e.: calculating odds ratios or prevalence estimates from in-depth interview and focus group data collected from convenience samples).

The main difficulty in shortlisting from title/abstract screening to full text screening was the variable quality of abstracts associated with the qualitative studies. If methods were not adequately discussed in the abstracts, we erred on the side of caution and over-included at this stage.

4. Was a software used for coding, as it helps in data organization and retrieval, especially if the volume of studies is large. More and more qualitative studies are using a software.

We used a combination of analysis methods in this review, including coding by hand in order to better engage the reviewers in dynamic discussion of the initial and axial themes, as well as organizing and managing the extracted data in both Excel and Atlas.ti. We found that using qualitative software during the analysis was less interactive, more static and more challenging to engage more than one reviewer at a time, particularly as we moved from initial to axial coding.

5. Results: The findings are well written but Quality of Care needs substantial enhancement as it is presently confined to perceptions related to staff attitudes. There is much more which can go here, very fundamental quality issues such as: 24/7 functionality of health centers, presence of female staff, availability of supplies and medicines.

Thank you for your feedback. We agree the points mentioned above (24/7 functionality of health centers, presence of female staff, availability of supplies and medicines) are certainly within the realm of quality of care. In this review, we focused on a user-centric model of facilitators and barriers to childbirth at a facility, and as such, focused this section of the review findings on perceived quality of care (i.e.: from the user perspective or from providers speaking about users’ perspectives).

In regards to the 24/7 availability of health centers, this topic is discussed previously in the “Resource availability and access” section. [“Travel at night or on weekends is especially difficult as there are fewer options and higher costs (20,31,37,40,42,43). Furthermore, health facilities may be closed or lack appropriate staffing to manage a delivery or complications at night (31,37). Lack of access to transportation, good roads, adequate funds, and communication systems also make organizing referrals for obstetric complications a time-consuming process (15,27,32,42,46).”]. Health service availability is also discussed in the first order theme in Table 1: “Barrier: Inaccessibility of transportation and facilities during off-hours (20,31,37,40,42,43).” as well as in
review finding #23 [“Barrier: Inaccessibility of transportation and facilities during off-hours
Travel at night or on weekends was considered particularly difficult as there are fewer public transportation options, women may be afraid of thieves and wild animals, and the price is higher. Even if women are able to arrange transportation during the off-hours, health facilities may be closed or lack the staffing to manage her delivery. (20,31,37,40,42,43)”]

In regards to the availability of supplies and medicines, this topic was not explicitly seen in the studies included in this review. This may be due to the user-centric focus that we employed, where users may not be as cognizant of supply shortages as a provider or health administrator might be, or perhaps the availability of supplies did not influence the user’s decision to deliver at a facility at an individual level. Indirect, point-of-care costs associated with facility-based deliveries are discussed in review finding #27 [“Barrier: Indirect and hidden costs associated with facility delivery
Even in settings where direct delivery costs were subsidized, families were expected to pay for transportation to the facility, drugs, medical supplies (i.e.: gloves, needles, gauze), blood for transfusions, laboratory services, food during the hospital stay, bribes to health providers, and laundry services. These additional costs often came as a surprise to women after they attended the facility, which may impact their future choice of delivery location. In addition to the extra point-of-care costs associated with facility birth, families experienced opportunity costs due to absence from work and domestic responsibilities. (14,17,18,20,22,23,28,29,32,40-42,44-46)” as well as in the “resource availability and access” section [“Women viewed costs outside of the direct cost for a delivery as hidden costs that were difficult to prepare for (14,17,18,20,22,23,28,29,32,40-42,44-46). Even in settings where direct delivery costs were subsidized, families were expected to pay for transportation to the facility, and other costs related to treatment at the facility (20,23,28,29,32,40,45,46).”]

The lack of presence of female staff as a barrier to the decision to deliver at a health facility was not seen in the studies included in this review.

6. Your Table 2 indicates High, Moderate and Low certainty of evidence. This needs to be brought into the narrative of the Findings section. As there are a number of findings reported here, there needs to be discrimination as which findings were reported more frequently and which were more anecdotal.

Thank you for your feedback. The CERQual methodology provides a rigorous process by which to assess the confidence in the review findings, so none of the findings reported either in the findings section, table 1, or table 2 are anecdotal. Rather, all review findings discussed in the text and corresponding tables were reported after several rounds of coding and analysis. One component of the CERQual tool is coherence, which refers to the assessment of the extent to which clear patterns could be identified from the data in the included primary studies. We are hesitant to bring the confidence assessments into the text of the findings section as this would represent a
duplication of the analysis already presented in table 2 (i.e.: see Colvin 2013 qualitative evidence synthesis using the CERQual approach that does not include confidence assessments in the text of the findings section).

7. The Introduction mentions a Conceptual Framework that would emerge form the Findings, however this missing. You have good material to come up with a Framework which would add to the strength of the paper.

We developed a multi-level life course conceptual framework from this review that can be found in Figure 2.

8. The discussion focuses primarily on fears related to undesirable birth practices, there is a lot more information that goes missing - such as travel, quality issues, costs etc. You need to bring in this information and then take a stance as to which were more recurring and crosscutting issues across countries. Also where data was thin, and could have biased results. Overall, I think this paper is worth investing in further improvements as it has the potential to make a very interesting read.

Thank you for sharing

Thank you for your feedback. Our primary aim in the discussion was to elucidate the newer or under-explored areas of research and practice that this review brought to light. While we recognize the importance of accessibility issues (i.e.: travel to facilities and prohibitive costs), these issues have been highlighted by research and practice in the past few decades since Thaddeus and Maine’s three delays model. We believe that one of the strengths of this review will be to help bring issues such as “improving intrapartum care” more firmly into the academic discourse on the seeking of facility-based delivery, and thus draw more attention to addressing these existing barriers to utilization of maternal health resources (and specifically to facility-based delivery). To this end, we chose to use the allot word limit to emphasize the dearth of research related to mistreatment of women and perceived low quality of care from the user perspective during the intrapartum period, with the hope that highlighting this important issue would generate more momentum around the topic. We added text to the limitations section to address where data might have been sparse.

Level of interest: An article of importance in its field
Quality of written English: Acceptable
Statistical review: No, the manuscript does not need to be seen by a statistician.
Declaration of competing interests: No competing interests