Incorporating qualitative research into guideline development: the way forward

Özge Tunçalp, MD PhD
Department of Reproductive Health and Research
Outline

- WHO guideline processes
  - Moving beyond effectiveness
- Development of CERQual
  - Assessing confidence in the evidence from reviews of qualitative research
- Implications
WHO Guideline Development Process

- Structured process
- In summary, it includes:

  (i) identification of priority questions and critical outcomes;
  (ii) retrieval of the evidence;
  (iii) assessment and synthesis of the evidence;
  (iv) formulation of recommendations;
  (iv) planning for dissemination, implementation, impact evaluation and updating.
What questions are decision makers asking?

Evidence of effectiveness is now seen as an important basis for healthcare decisions

Is the intervention effective and safe?

But decision makers are also asking other questions

Will it be acceptable to patients and others?

Will it be feasible to implement?
Asking additional questions in WHO guidelines

All of these questions were asked in 3 WHO guidelines:

1. Task-shifting for MNH healthcare (2012)
3. Antenatal care (upcoming)
Using the DECIDE framework

- For these three guidelines, WHO chose to use the DECIDE framework
- Helps decision makers consider a range of relevant criteria when making decisions, including:
  - Benefits + harms
  - Acceptability
  - Feasibility
  - Resource use
  - Equity
DECIDE wants to move away from this type of decision making....
...to a more systematic and transparent assessment of relevant criteria
...to a more systematic and transparent assessment of relevant criteria
...to a more systematic and transparent assessment of relevant criteria
Evidence about acceptability and feasibility

- Is the intervention effective and safe?
  - Systematic reviews of controlled studies

- Is the intervention acceptable?
  - Systematic reviews of qualitative research

- Is the intervention feasible?
Assessing confidence in the evidence

- When presenting evidence about *effectiveness* to decision makers, WHO uses GRADE

- Here, our confidence in findings from effectiveness reviews can be downgraded when:
  - The studies are poorly designed or poorly run
  - The studies focus on slightly different topics than the review question
  - There are few studies or the studies have insufficient data
  - The results vary across studies

- But what about evidence from reviews of qualitative research?
GRADE CERQual

Confidence in the Evidence from Reviews of Qualitative research
Why did we develop CERQual?

- Systematic reviews of qualitative research (also called qualitative evidence syntheses) becoming increasingly common

- Also increasingly being used in guideline or policy development processes

- Users need methods to assess how much confidence to place in findings from these reviews

- Users are likely to make these judgements anyway – it may be helpful to provide a systematic and transparent way for doing this
In developing CERQual, we needed an approach that:

- Could be applied to the typical types of qualitative study approaches (e.g. ethnography,) and data (e.g. from interviews, focus groups etc.)

- Was easy to use

- Allowed judgements to be reported transparently

- Allowed the judgements to be understood, including by users without an in-depth understanding of qualitative methods
Developing the CERQual approach

- Started in 2010 to support the use of qualitative evidence syntheses in a new WHO guideline (OptimizeMNH)
- *First version*: two components
- Broad consultation with wide group of stakeholders
- *Second version*: four components
Relationship to GRADE

- CERQual is part of the GRADE Working Group

- CERQual shares the same aim as the GRADE tool used to assess the certainty of evidence of *effectiveness*

- However, CERQual is grounded in the principles of qualitative research
CERQual is **not** a tool for:

- Assessing how well an individual qualitative study was conducted
- Assessing how well a systematic review of qualitative studies was conducted
- Assessing confidence in ‘narrative’ or ‘qualitative’ summaries of the effectiveness of an intervention, where meta-analysis is not possible
What **does** the CERQual approach do?

- CERQual aims to transparently assess and describe how much confidence to place in findings from qualitative evidence syntheses.
CERQual is applied to **individual synthesis findings**

- In the context of a qualitative evidence synthesis, a finding is:
  ...an analytic output that describes a phenomenon or an aspect of a phenomenon

- Findings from qualitative evidence syntheses can be presented as:
  - themes, categories or theories
  - both descriptive or more interpretive findings
What do we mean by ’confidence in the evidence’?

An assessment of the extent to which a review finding is a reasonable representation of the phenomenon of interest

• i.e. the phenomenon of interest is unlikely to be substantially different from the review finding
The CERQual approach

Overall aim of the approach:
To assess how much confidence we have in the evidence for the review finding

This is based on an assessment of

- METHODOLOGICAL LIMITATIONS of the individual studies contributing to the review finding
- RELEVANCE to the review question of the individual studies contributing to the review finding
- COHERENCE of the review finding
- ADEQUACY OF DATA contributing to the review finding
CERQual in the literature

Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: qualitative evidence synthesis (Review)

Glenton C, Colvin CJ, Carlsen B, Swartz A, Lewin S, Noyes J, Rashidian A

GUIDELINES AND GUIDANCE

Using Qualitative Evidence in Decision Making for Health and Social Interventions: An Approach to Assess Confidence in Findings from Qualitative Evidence Syntheses (GRADE-CERQual)

Simon Lewin1,2*, Claire Glenton1, Heather Munthe-Kaas3, Benedicte Carlsen4, Christopher J. Colvin5, Metin Gülmezoglu6, Jane Noyes7, Andrew Booth8, Ruth Garside9, Arash Rashidian10,11

RESEARCH ARTICLE

The Mistreatment of Women during Childbirth in Health Facilities Globally: A Mixed-Methods Systematic Review

Meghan A. Bohren1,2*, Joshua P. Vogel2, Erin C. Hunter2, Olha Lutsiv4, Suprita K. Makh5, João Paulo Souza6, Carolina Aguiar7, Fernando Saraiva Coneglian6, Alex Luiz Araújo Diniz6, Özge Tunçalp2, Dena Javadi3, Olufemi T. Oladapo2, Rajat Khosla2, Michelle J. Hindin1,2, A. Metin Gülmezoglu2
Component 1: Methodological limitations

The extent to which there are *problems in the design or conduct* of the primary studies supporting a review finding
Concerns about methodological limitations

We are less confident that the finding reflects the phenomenon of interest when:

- the primary studies underlying a review finding are shown to have **problems in the way they were designed or conducted**

- A quality appraisal tool for qualitative studies should be used to make this assessment
  - Typically includes appraisals of how the participants and settings were selected, how data was collected and analysed, researcher reflexivity etc

- Currently no widespread agreement about the best tool – research agenda in place
Component 2: Relevance

The extent to which the body of evidence from the primary studies supporting a review finding is applicable to the context specified in the review question.
Concerns about relevance

We are less confident that the finding reflects the phenomenon of interest when:

- the contexts of the primary studies underlying a review finding are **substantively different from the context** of the review question

• For instance, where the relationship between the studies and the review question context is:
  - **Partial** - studies only represent part of the phenomenon of interest
  - **Indirect** - studies do not directly reflect the phenomenon of interest
  - **Unsure / uncertain**
The extent to which the finding is *well grounded in data* from these primary studies and *provides a convincing explanation* for the patterns found in these data.
Concerns about coherence

We are less confident that the finding reflects the phenomenon of interest when:

There are variations or exceptions in the data and there is no convincing explanation for these variations or exceptions.
Component 4: Adequacy of data

The degree of *richness* and *quantity of data* supporting a review finding
Concerns about adequacy of data

We are less confident that the finding reflects the phenomenon of interest when:

- the data underlying a review finding is not sufficiently rich or only comes from a small number of studies or participants

- Review authors need to make a judgement on what constitutes data that is not sufficiently rich or too small a number in the context of a specific review finding
After assessing each of the separate components, we make an overall judgement of the confidence in each review finding.
Confidence can be assessed as high, moderate, low or very low

- **High confidence**: It is highly likely that the review finding is a reasonable representation of the phenomenon of interest.

- **Moderate confidence**: It is likely that the review finding is a reasonable representation of the phenomenon of interest.

- **Low confidence**: It is possible that the review finding is a reasonable representation of the phenomenon of interest.

- **Very low confidence**: It is not clear whether the review finding is a reasonable representation of the phenomenon of interest.
Summary of qualitative findings tables

- The CERQual assessment must be described transparently in a Summary of Qualitative Findings (SoQF) table

- Summarises the key findings and the confidence in the evidence for each finding, and provides an explanation of the assessment of the confidence in the qualitative evidence

<table>
<thead>
<tr>
<th>Review Finding</th>
<th>Contributing Studies</th>
<th>Confidence in the Evidence</th>
<th>Explanation of Confidence in the Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Physical abuse</strong></td>
<td>[6, 9, 10, 13, 21, 61, 67, 66, 73, 75, 77, 80, 84, 86, 87, 91, 96, 97]</td>
<td>High</td>
<td>18 studies with minor to significant methodological limitations. Thick data from 11 countries across all geographical regions, but predominantly sub-Saharan Africa. High coherence.</td>
</tr>
<tr>
<td><strong>Verbal abuse</strong></td>
<td>[13, 58, 59, 63, 68, 77]</td>
<td>Moderate</td>
<td>6 studies with minor to significant methodological limitations. Adequate data from 5 countries, predominantly middle- and high-income countries. High coherence.</td>
</tr>
</tbody>
</table>

**Use of force:** Women across the world reported experiencing physical force by health providers during childbirth. In some cases, women reported specific acts of violence committed against them during childbirth, but women often referred to these experiences in a general sense and alluded to beatings, aggression, physical abuse, a rough touch, and use of extreme force. Pinching, hitting, and slapping, either with an open hand or an instrument, were the most commonly reported specific acts of physical violence.

**Threats and blaming:** Some women were threatened with poor quality of care or poor outcomes for their babies as a result of their behavior during childbirth. This included threats of a beating if the woman did not comply with a health worker’s request and threats of withholding health services. Other women were blamed for their baby’s or their own poor health outcomes.
CERQual has the potential to:

- **Facilitate reflection** on findings from qualitative evidence syntheses
- Make more **explicit**:
  - where there are **gaps or insufficient evidence** in relation to a review question
  - where **only poor quality studies** are available
- Prompt the generation of **new explanations or concepts to explain patterns** in findings

**Applying CERQual involves judgements.**

CERQual attempts to make these judgements transparent and increase the contribution of qualitative research to decision making.
From reviews to recommendations

Reviews completed → GRADE, SOF & CERQual, SOQF → DECIDE Framework → Draft recommendation

Evidence-to-decision framework
## Example

### Recommendations for vacuum aspiration for induced abortion

<table>
<thead>
<tr>
<th>Health worker</th>
<th>Recommendation</th>
<th>Justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Midwives</td>
<td>Recommended</td>
<td>There is evidence for the safety and effectiveness (moderate certainty) and for women’s satisfaction with the overall abortion experience (low certainty). This task is recognized as a core competency in midwifery. Women often consider care received from midwives as more supportive (moderate confidence). The option has been shown to be feasible, including in low-resource settings.</td>
</tr>
<tr>
<td>Nurses</td>
<td>Recommended</td>
<td>There is evidence for the safety and effectiveness (low certainty) and for women’s satisfaction with this option (low certainty). Women often consider care received from nurses as more supportive (moderate confidence). The option is feasible and may decrease inequities by extending safe abortion care to underserved populations.</td>
</tr>
<tr>
<td>Auxillary nurses (AN) and auxillary nurse midwives (ANM)</td>
<td>Recommended in specific circumstances</td>
<td>Although there was insufficient direct research evidence for the effectiveness of this option, the benefits outweigh any possible harms. The option has also been shown to be feasible, including at scale in low-resource settings, and has the potential to decrease inequities by extending safe abortion care to rural and underserved populations.</td>
</tr>
</tbody>
</table>

We recommend this option in contexts where established mechanisms to include ANMs/ANs in providing basic emergency obstetric care or post-abortion care already exist.
What are the implications that we have observed from the experiences of including a broader evidence base in WHO guidelines?
The inclusion of a wider range of evidence requires:

- More resources
- A broader set of skills within one team
- Brings a multi-dimensional and in-depth look
It can facilitate the uptake of the evidence and guidelines

- Some guideline panel members sceptical to trial focus - dislike neglect of “programmatic experience” and/or "women/user-centred evidence"

- Incorporating qualitative evidence allows for a stronger foundation on "implementation implications" for each question
It can reduce the use of anecdotal evidence

- Especially regarding acceptability, feasibility and related implementation implications
In conclusion

- Asking the **right questions** and using the **appropriate evidence** to support the decision-making process
  - DECIDE frameworks allow decision makers to assess different types of evidence systematically in a transparent way

- Syntheses of qualitative evidence can offer **useful information about acceptability and feasibility** – but we need to indicate to decision makers how confident we are in this evidence
  - CERQual approach

- WHO plans to continue to include qualitative evidence
  - New chapter on WHO handbook
Acknowledgements

- GRADE-CERQual
- DECIDE
- Cochrane
- World Health Organization
Thank you!

For more information,

Follow us on Twitter @HRPresearch
@CERQualNet

Website who.int/reproductivehealth
www.cerqual.org

tuncalpo@who.int
@otuncalp