University of Nottingham
UK | CHINA | MALAYSIA

JOANNA BRIGGS COLLABORATION
The Nottingham Centre for Evidence-Based Healthcare
A Joanna Briggs Institute Centre of Excellence

Establishing Confidence in the Findings from Qualitative Evidence Synthesis: Reflections on CONQUAL and CERQUAL

Dr Catrin Evans

Objectives

1. Using an example from a recent qualitative evidence synthesis, to identify some conceptual and practical challenges associated with CONQUAL and CERQUAL

2. To compare the two approaches and to discuss their respective strengths and weaknesses

3. To suggest an agenda for future research

University of Nottingham
UK | CHINA | MALAYSIA

JOANNA BRIGGS COLLABORATION
The Nottingham Centre for Evidence-Based Healthcare
A Joanna Briggs Institute Centre of Excellence

Example Review: Improving care for women and girls who have experienced Female Genital Mutilation/Cutting (FGM/C)

Exemplar Review: Female Genital Mutilation/Cutting (FGM/C)

- Funded by NIHR
- Aimed to explore the experiences of accessing and receiving FGM/C-related healthcare across the life course for women and girls who have undergone FGM/C (in OECD countries)
- Qualitative evidence synthesis, using a thematic synthesis approach (& NVivo software)
- Utilised the JBI QARI critical appraisal tool
- Applied CERQUAL to assess confidence in the review findings (but also considered CONQUAL for this presentation)
- Large review – included 57 studies
- Synthesis included 17 descriptive themes, 5 analytical themes & a conceptual model outlining culturally safe care

University of Nottingham
UK | CHINA | MALAYSIA

JOANNA BRIGGS COLLABORATION
The Nottingham Centre for Evidence-Based Healthcare
A Joanna Briggs Institute Centre of Excellence

Assessing Confidence in Review Findings

JBI: CONQUAL

- Confidence: "the belief or trust that a person can place in the results of research" (Munn et al, 2014:3)
- Uses two concepts: dependability and credibility
- Dependability (akin to reliability): evaluates whether the process of research is logical, traceable and clearly documented, particularly on the methods chosen and the decisions made by the researchers
  - Relates to methodological quality of included studies
- Credibility (akin to internal validity): evaluates the congruence between the author's interpretation and the supporting data – considers whether the findings authentically and comprehensively reflect the phenomenon of interest
  - Relates to "concept-indicator fit" (Seale, 1999) - a global evaluation of the 'fit' between the primary data and the reviewer's interpretations

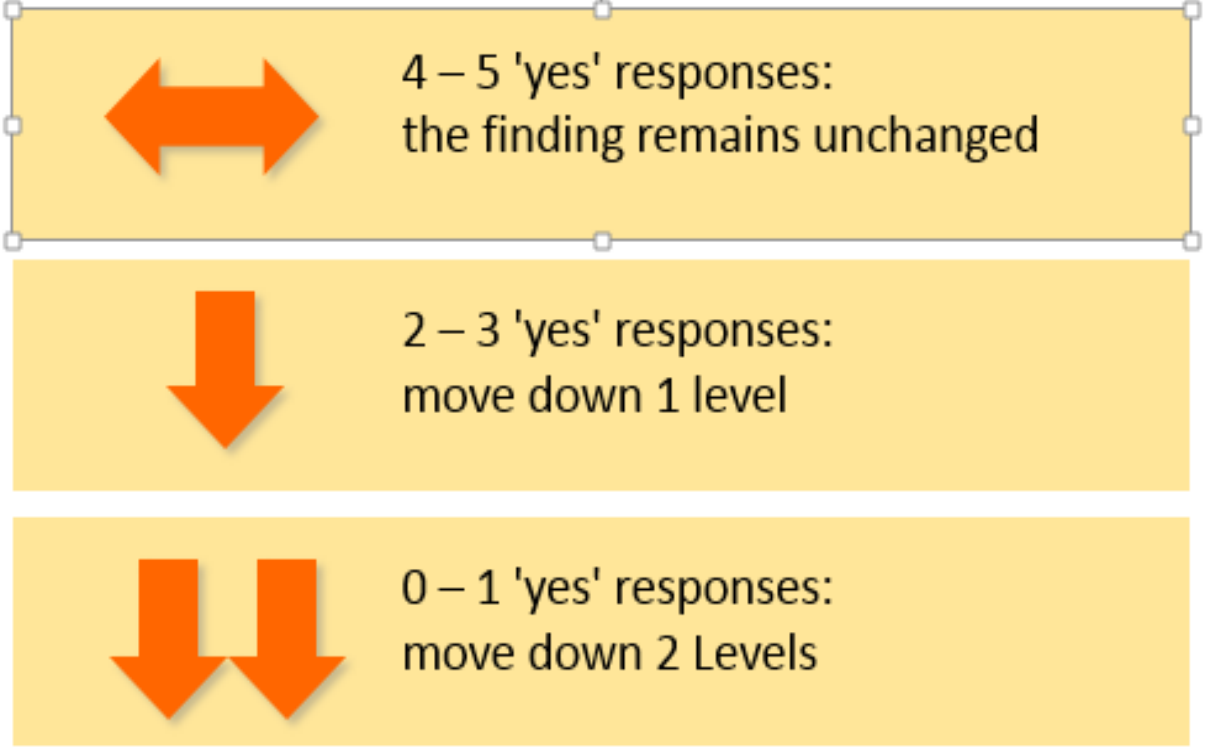
The ConQual Approach: Dependability (Methodological Quality)

Measurement

Measured by asking the following questions:

1. Is there congruity between the research methodology and the research question or objectives?
2. Is there congruity between the research methodology and the methods used to collect data?
3. Is there congruity between the research methodology and the representation and analysis of data?
4. Is there a statement locating the researcher culturally or theoretically?
5. Is the influence of the researcher on the research, and vice-versa, addressed?

Ranking system:



The ConQual Approach: Credibility

Measurement

Assign a level of credibility to the findings:

Unequivocal (findings accompanied by an illustration that is beyond reasonable doubt and; therefore not open to challenge)

Credible (findings accompanied by an illustration lacking clear association with it and therefore open to challenge)

Unsupported (findings are not supported by the data, or with no illustration)

Ranking

The synthesized findings contains only unequivocal findings
No change

Mix of unequivocal/credible findings
Downgrade one level (-1)

All credible findings
Downgrade two levels (-2)

Mix of credible/unsupported findings
Downgrade three levels (-3)

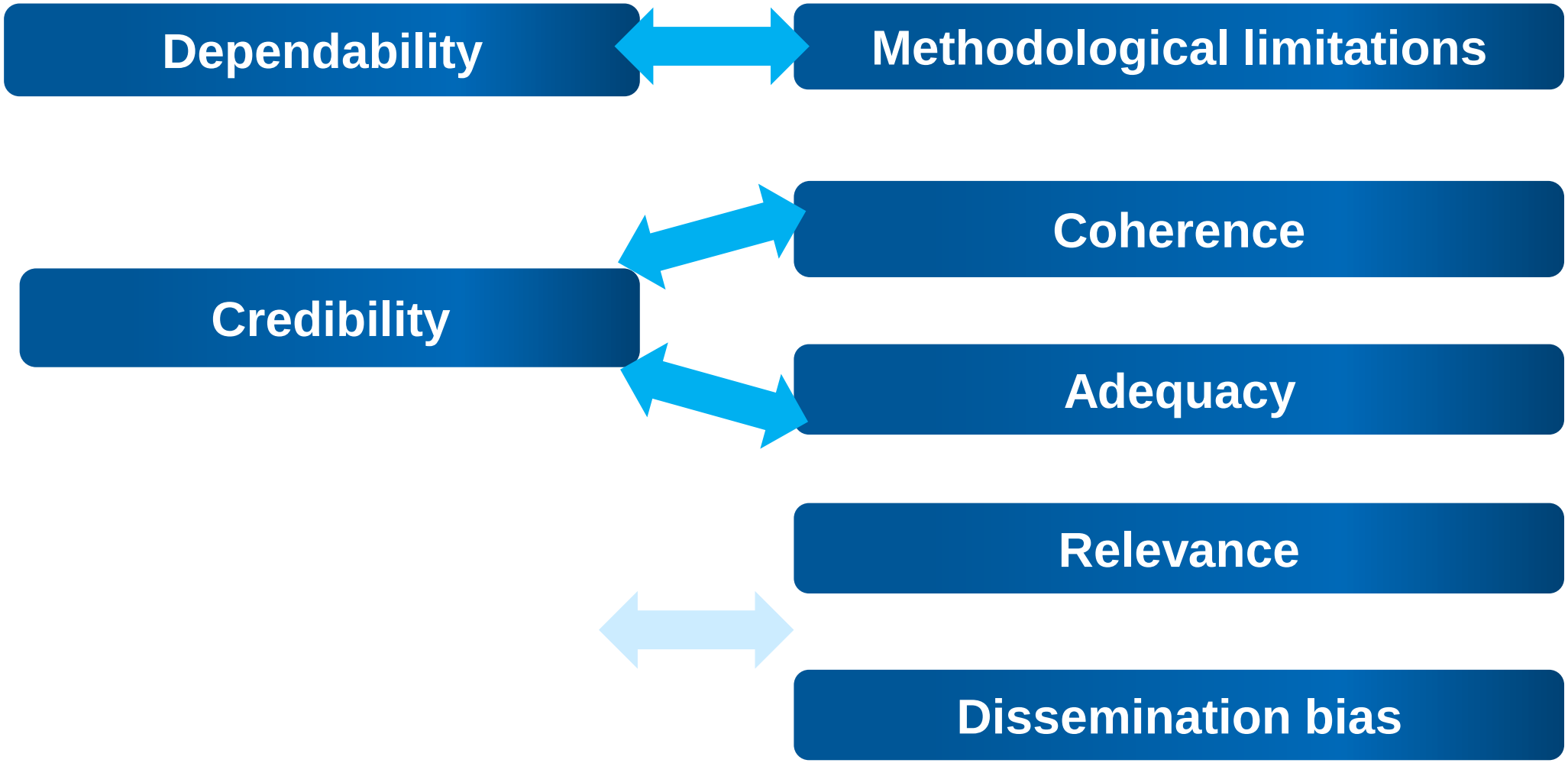
Unsupported findings
Downgrade four levels (-4)

GRADE-CERQUAL

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

Component	Definitions	Assessment of Confidence
Methodological limitations	The extent to which there are problems in the design of conduct of the primary studies that contributed to the review	**Assessment in 4 Domains** • No concerns • Minor concerns • Moderate concerns • Serious concerns **Summary of Findings Table** • High confidence • Moderate confidence • Low confidence • Very low confidence
Relevance	The extent to which the body of evidence from the primary studies supporting a review finding is applicable to the context (perspective, population, phenomenon of interest or setting) specified in the review question	
Coherence	The extent to which the review finding is well grounded in data from the contributing primary studies and provides a convincing explanation for the patterns found in the data	
Adequacy	An overall determination of the degree of richness and quantity of data supporting a review finding	
Dissemination bias	Under development	

CONQUAL – CERQUAL Comparison



Methodological Limitations & Dependability: FGM/C Review

CONQUAL

- Looking just at dependability - out of 57 studies, only 21% (n=12) would have been considered ‘good’ & not downgraded.

The majority (58%, n=33) would have been downgraded one level and 21% (n=12) would have been downgraded two levels.

Methodological Limitations & Dependability: FGM/C Review

CERQUAL

- Looking at methodological quality/limitations as an overall assessment, we used a weighting system to assign a quality band to different studies as: high, medium or low
- High was any study that scored >7 on the JBI QARI, medium was 5-7 and low was <5
- Using this approach, 52% (n=30) were considered high quality, 37% (n=21) were medium quality and 11% (n=6) were low quality

Due to the mix of studies within each finding, all review findings were downgraded from ‘no concerns’ to ‘moderate concerns’

Challenges: Assessing Dependability or Methodological Limitations

Using checklists to assess quality – are we conflating reporting with quality?

- Papers which may score badly on a quality checklist (especially when written in the context of other disciplinary traditions) may contain important insights (Ray Pawson, 1998: “Digging for nuggets”)
- Papers which score well, may contain ‘thin’ uninteresting findings
- It remains unclear to what extent poor quality papers would influence the findings of a review, especially when there are many studies (sensitivity analysis)

FGM/C Review……

- Only five studies (9%) reported fully on both the reflexivity questions (Qs. 6&7) and 34 studies (60%) did not mention this aspect at all (this had a major impact on the CONQUAL dependability rating)
- Ten studies (18%) did not report having obtained ethical approval (yet seemed to have been conducted in an entirely appropriate manner)

Challenges: Assessing Dependability or Methodological Limitations

CONQUAL

- Once ranked, the 'grading' of each paper is fixed, so relatively easy to apply across the different review findings

CERQUAL

- Requires a new analysis of the potential impact of methodological flaws for each individual review finding. Challenging to undertake such a detailed & nuanced analysis when there are large numbers of studies

Making a Judgement on the Rating

- The FGM/C review findings each had between 19-46 underpinning studies – with varying methodological quality 'rankings'
- When each review finding has a large number of studies, how should the final rating be determined? Should the rating reflect the majority? Or average? Or the most methodologically strong (even if small numbers)? - needs in-depth discussion and justification on a case by case basis

CONQUAL: Assessing Credibility

- FGM/C Review – the majority of findings from the primary studies were classified as either unequivocal or credible. But some were 'unsupported'
- **CONQUAL** – all 17 review findings were a mix mainly of unequivocal and credible (with a small number of unsupported)
- It was hard to make a judgement due to this mix of credibility levels within the findings
- Detailed guidance was lacking

As per current CONQUAL guidance, all review findings would have been downgraded at least one or two levels

CERQUAL: Assessing Coherence

- **Coherence:** The extent to which a review finding is well grounded in data from the contributing primary studies and provides a convincing explanation for the patterns found in the data, i.e. how clear and cogent is the fit between the data from the primary studies and the review finding?
- Considers consistency (and difference) across studies
- Re-examined the underpinning data for each theme. Assessed clarity of concepts, consistency and the extent to which they applied to similar or different groups or contexts or manifested under particular conditions

7/17 review findings judged to have no concerns, 3/17 had minor concerns and 7/17 had 'moderate' concerns

Importance of heterogeneity & the disconfirming case

"Rather than disturbing rigor, data incoherence may actually contribute to exciting analysis leading to new knowledge."
(Malterud, 2018:6)

CERQUAL: Assessing Adequacy

- **Adequacy**: An overall determination of the degree of **richness** and **quantity** of data supporting a review finding (similar to 'theoretical saturation')
- Utilised an assessment previously described by Popay et al (1998) to classify study **design** and **findings** (overall) in terms of richness ('thick' or 'thin')
- Examined the number of studies reporting a theme within particular contexts or within particular groups (* but careful not to apply a quantitative logic here *)

3/17 findings had no concerns, 13/17 had minor concerns, 1/17 had moderate concerns

Richness	Operational Definition
Thick papers	• **Offer greater explanatory insights into the outcome of interest** • Provide a clear account of the process by which the findings were produced—including the sample, its selection and its size, with any limitations or bias noted—along with clear methods of analysis • **Present a developed and plausible interpretation** of the analysis based on the data presented.
Thin papers	• **Offer only limited insights** • Lack a clear account of the process by which the findings were produced • **Present an underdeveloped and weak interpretation** of the analysis based on the data presented

Assessing Relevance

- **Relevance**: The extent to which the body of evidence from the primary studies supporting a review finding is applicable to the context (perspective, population, phenomenon of interest or setting) specified in the review question
- Re-examined the underpinning data for each review finding to assess the degree to which the findings applied to the contexts, groups and issues specified in the review question (examining application to similar or different groups or contexts or types of FGM/C)

7/17 review findings assessed as no concerns; 7/17 assessed as minor concerns; 3/17 assessed as moderate concerns

Overall Assessment of Confidence

FGM/C Review: 17 Review Findings	
CONQUAL	• All 17 rated ‘moderate’
CERQUAL	• 10 rated ‘high’ • 7 rated ‘moderate’

University of Nottingham
UK | CHINA | MALAYSIA

JOANNA BRIGGS COLLABORATION
The Nottingham Centre for Evidence-Based Healthcare
A Joanna Briggs Institute Centre of Excellence

Final Reflections

CONQUAL/CERQUAL: Main Differences

Differences

- Different approaches to assessing methodological limitations/dependability
- CONQUAL assesses 2 dimensions, CERQUAL assesses 4
- CONQUAL is quicker and easier to apply!
- But – there is less room for nuance or judgement within the CONQUAL domains
- CONQUAL relies on use of the JBI QARI tool

CONQUAL – way forward

- Needs more detailed guidance and some worked examples

Practical Issues: Both Approaches

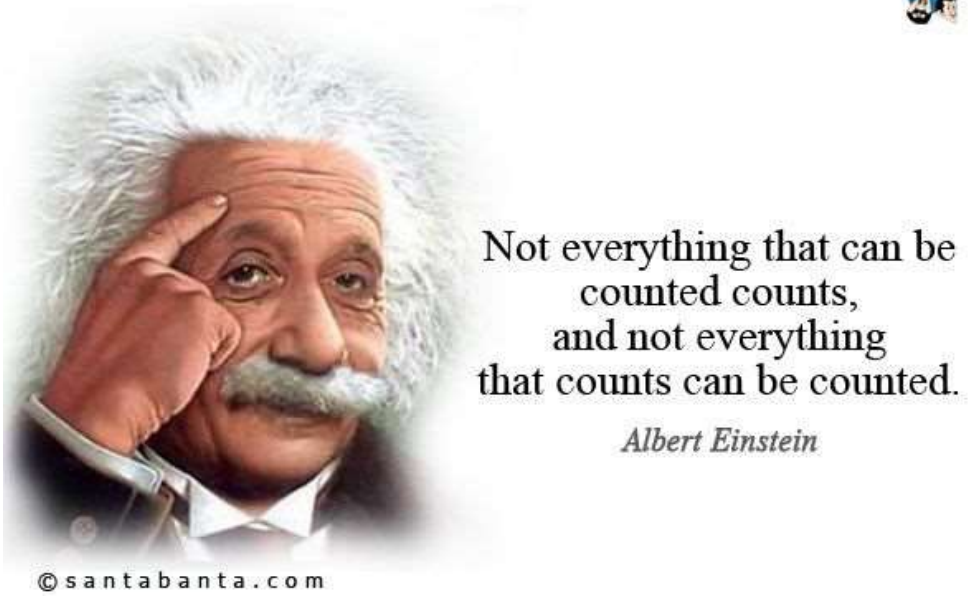
- Difficulty with managing large numbers of studies
- Extra time/resource implications for completion of the review
- Software is not yet set up to support these assessments

Conceptual Issues: Both Approaches

Confidence assessment – essential for guideline development & decision making contexts

- But – how to do this without becoming too procedural and risk losing the focus on creativity of interpretation inherent within qualitative research?

- Could risk becoming quantitative in logic and formulaic (too much focus on 'scores' and aggregate sums)



"When procedure over-rules discovery, analysis suffers, resulting in trivial reports that are boring to read, though conducted according to methodological standards" (Kvale 1996)

Future Research Agenda

- More research comparing the two approaches across different teams and different reviews
- Research on the relative impact of different quality assessment tools on confidence ratings
- Research on how user friendly and understandable confidence assessments are for guideline developers – what impact are they having? (how, where and why are they being used?), how do guideline developers understand the complexities underpinning the confidence ratings?
- How to assess confidence in mixed methods reviews?
- How to assess confidence in more theoretical review outputs? (e.g. conceptual models or logic models)

References

Booth A, Lewin S, Glenton C, Munthe-Kaas H, Toews I, Noyes J, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings–paper 7: understanding the potential impacts of dissemination bias. Implement Sci. 2018;13(1):12.

Colvin CJ, Garside R, Wainwright M, Munthe-Kaas H, Glenton C, Bohren MA, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 4: how to assess coherence. Implement Sci. 2018;13(1):13.

Glenton C, Carlsen B, Lewin S, Munthe-Kaas H, Colvin CJ, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 5: how to assess adequacy of data. Implement Sci. 2018;13(1):14.

Kvale S. InterViews: An Introduction to Qualitative Research Interviewing. Thousand Oaks, California: Sage Publications; 1996.

Lewin S, Bohren M, Rashidian A, Munthe-Kaas H, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: how to make an overall CERQual assessment of confidence and create a Summary of Qualitative Findings table. Implement Sci. 2018;13(1):10.

Lewin S, Booth A, Glenton C, Munthe-Kaas H, Rashidian A, Wainwright M, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. Implement Sci. 2018;13(1):2.

Lewin S, Glenton C, Munthe-Kaas H, Carlsen B, Colvin C, Gulmezoglu M, et al. Using qualitative evidence in decision making for health and social interventions: an approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQUAL). PLoS Med. 2015;12(10):e1001895.

Malterud K. The impact of evidence based medicine on qualitative metasynthesis: benefits to be harvested and warnings to be given. Qual Health Res. 2018.

Munn Z, Porritt K, Lockwood C, Aromataris E, Pearson A. Establishing confidence in the output of qualitative research synthesis: the ConQual approach. BMC Med Res Meth. 2014;14:108.

Munthe-Kaas H, Bohren MA, Glenton C, Lewin S, Noyes J, Tunçalp Ö, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 3: how to assess methodological limitations. Implement Sci. 2018;13(1):9.

Noyes J, Booth A, Lewin S, Carlsen B, Glenton C, Colvin CJ, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings–paper 6: how to assess relevance of the data. Implement Sci. 2018;13(1):4.

Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. Qual Health Res. 1998;8.

Pawson R. Digging for nuggets: how 'bad' research can yield 'good' evidence. International Journal of Social Research Methodology. 2006;9(2):127-42.

Seale C. The Quality of Qualitative Research. London: Sage Publications; 1999.

Thorne S. Metasynthetic madness: what kind of monster have we created? Qual Health Res. 2017;27(1):3-12.

Toye F, Seers K, Allcock N, Briggs M, Carr E, Barker K. Meta-ethnography 25 years on: challenges and insights for synthesisng a large number of qualitative studies BMC Med Res Meth. 2014;14:80.

Toye F, Seers K, Hannink E, Barker K. A mega ethnography of eleven qualitative evidence syntheses exploring the experience of living with chronic non-malignant pain. BMC Med Res Methodol. 2017;17(116).

THANK YOU

ANY QUESTIONS?