

Chapter 6 – Critical appraisal of qualitative research

Key points

- Critical appraisal of qualitative studies is an essential step within a Cochrane Intervention review that incorporates qualitative evidence.
- The overarching goal of critical appraisal in the context of including qualitative research in a Cochrane Intervention Review is to assess whether the studies actually address questions under meaning, process and context in relation to the intervention and outcomes under review.
- Review teams should use a critical appraisal instrument that is underpinned by a multi-dimensional concept of quality in research and hence includes items to assess quality according to several domains including quality of reporting, methodological rigour and conceptual depth and breadth.
- Critical appraisal involves (i) filtering against minimum criteria, involving adequacy of reporting detail on the data sampling, -collection and-analysis, (ii) technical rigour of the study elements indicating methodological soundness and (iii) paradigmatic sufficiency, referring to researchers' responsiveness to data and theoretical consistency.
- When choosing an appraisal instrument a Review teams should consider the available expertise in qualitative research within the team and should ensure that the critical appraisal instrument they choose is appropriate given the review question and the type of studies to be included.
- Reviewers need to clarify how the outcome of their critical appraisal exercise is used with respect to the presentation of their findings. The inclusion of a sensitivity analysis is recommended to evaluate the magnitude of methodological flaws or the extent to which it has a small rather than a big impact on the findings and conclusions.

Introduction

Considerable debate exists on whether or not concepts such as validity and reliability apply to qualitative research and if so how they could be assessed. Some researchers have stated that qualitative research should establish validity, reliability and objectivity. Others plead for an adjustment of these concepts to better fit the qualitative research design. As a consequence, critical appraisal instruments might differ in the criteria they list to complete a critical appraisal exercise. Some researchers consider appraisal instruments a tool that can be

utilized as part of the exploration and interpretation process in qualitative research (Popay et al, 1998; Spencer, 2003). Edwards et al (2002) describes the use of a “signal to noise” approach, where a balance is sought between the methodological flaws of a study and the relevance of insights and findings it adds to the overall synthesis. Other researchers do not acknowledge the value of critical appraisal of qualitative research, stating that it stifles creativity (Dixon-Woods, 2004). While recognising that all these views have some basis for consideration certain approaches succeed in positioning the qualitative research enterprise as one that can produce a valid, reliable and objective contribution to evidence synthesis. It is these that may therefore have more potential to be generally accepted within the context of producing Cochrane Intervention Reviews. The Cochrane Collaboration recommends a specific tool for assessing the risk of bias in each included study in an intervention review, a process that is facilitated through the use of appraisal instruments addressing the specific features of the study design and focusing on the extent to which results of included studies should be believed. This suggest that in assessing the methodological quality of qualitative studies the core criterion to be evaluated is researcher bias. Believability in this context refers to the ability and efforts of the researcher to make his or her influence and assumptions clear and to provide accurate information on the extent to which the findings of a research report hold true. However, it is the actual audit trail provided by researchers that allows for an in-depth evaluation of a study. Most existing appraisal instruments use broader criteria that account for reporting issues as well. We suggest that these issues should be part of the appraisal exercise. Currently, there are four possibilities to make use of qualitative research in the context of Cochrane Intervention reviews:

1. The use of qualitative research to define and refine review questions a Cochrane Review (**informing reviews**).
2. The use of qualitative research identified whilst looking for evidence of effectiveness (**enhancing reviews**).
3. The use of findings derived from a specific search for qualitative evidence that addresses questions related to an effectiveness review (**extending reviews**).
4. Conducting a qualitative evidence synthesis to address questions other than effectiveness (**supplementing reviews**).

The latter use (Supplementing) is beyond the scope of current Cochrane Collaboration policy (Noyes et al, 2008). Stand alone qualitative reviews that supplement Cochrane Intervention reviews need to be conducted and published outside of the Cochrane context.

Critical appraisal applies to all of the above possibilities.

Reviewers should bear in mind that narratives used in reports of quantitative research cannot be considered qualitative findings if they do not use a qualitative method of datacollection and –analysis. Therefore, critical appraisal based on

instruments developed to assess qualitative studies is not applicable to reports that do not meet the criteria of being a 'qualitative study'..

This chapter breaks down in four sections. **Section 1** addresses translated versions of core criteria such as validity, reliability, generalisability and objectivity of qualitative studies. **Section 2** presents an overview of different stages involved in quality assessment. **Section 3** guides the researcher through some of the instruments and frameworks developed to facilitate critical appraisal and **section 4** formulates suggestions on how the outcome of an appraisal of qualitative studies can be used or reported in a systematic review.

Section 1: Core criteria for quality assessment

Critical appraisal is “the process of systematically examining research evidence to assess its validity, results and relevance before using it to inform a decision” (Hill & Spittlehouse, 2003). Instruments developed to support quality appraisal usually share some basic criteria for the assessment of qualitative research. These include the need for research to have been conducted ethically, the consideration of relevance to inform practice or policy, the use of appropriate and rigorous methods and the clarity and coherence of reporting (Cohen & Crabtree, 2008). Other criteria are contested, such as the importance of addressing reliability, validity, and objectivity, strongly related to researcher bias. Qualitative research as a scientific process needs to be “rigorous” and “trustworthy” to be considered as a valuable component of Cochrane systematic review. Therefore an evaluation using such criteria is essential. Nevertheless we should acknowledge that the meaning assigned to these words may differ in the context of qualitative and quantitative research designs (Spencer et al, 2003).

Does translation of terminology compromise critical appraisal?

The concepts used in table 1 are based on Lincoln and Guba’s (1985) translation of criteria to evaluate the trustworthiness of findings. Acknowledging the difference in terminology does not obviate the rationale or process for critical appraisal. There might be good congruence between the intent of meanings relevant to key aspects of establishing study criteria, as demonstrated in table 1.

Table 1: Criteria to critically appraise findings from qualitative research

Aspect	Qualitative Term	Quantitative Term
Truth value	Credibility	Internal Validity
Applicability	Transferability	External Validity or generalisability
Consistency	Dependability	Reliability

Neutrality	Confirmability	Objectivity
------------	----------------	-------------

This scheme outlines some of the core elements to be considered in an assessment of the quality of qualitative research. However, the concept of confirmability might not be applicable to approaches inspired by phenomenology or critical paradigms in which the researcher's experience becomes part of the data (Morse, 2002). The choice of critical appraisal instruments should preferably be inspired by those offering a multi-dimensional concept of quality in research. Apart from methodological rigour, that would also include quality of reporting and conceptual depth and bread.

What indications are we looking for in an original research paper?

There are a variety of evaluation techniques that authors might have included in their original reports, that facilitate assessment by a reviewer and that are applicable to a broad range of different approaches in qualitative research. However, it should be stated that some of the techniques listed only apply for a specified set of qualitative research designs.

- **Assessing Credibility:** *Credibility* evaluates whether or not the representation of data fits the views of the participants studied, whether the findings hold true.
Evaluation techniques include: having outside auditors or participants validate findings (member checks), peer debriefing, attention to negative cases, independent analysis of data by more than one researcher, verbatim quotes, persistent observation etc.
- **Assessing Transferability:** *Transferability* evaluates whether research findings are transferable to other specific settings.
Evaluation techniques include: providing details of the study participants to enable readers to evaluate for which target groups the study provides valuable information, providing contextual background information, demographics, the provision of thick description about both the sending and the receiving context etc.
- **Assessing Dependability:** *Dependability* evaluates whether the process of research is logical, traceable and clearly documented, particularly on the methods chosen and the decisions made by the researchers.
Evaluation techniques include: peer review, debriefing, audit trails, triangulation in the context of the use of different methodological approaches to look at the topic of research, reflexivity to keep a self-critical account of the research process, calculation of inter-rater agreements etc.

- **Assessing Confirmability**: *Confirmability* evaluates the extent to which findings are qualitatively confirmable through the analysis being grounded in the data and through examination of the audit trail. Evaluation techniques include: assessing the effects of the researcher during all steps of the research process, reflexivity, providing background information on the researcher's background, education, perspective, school of thought etc.

The criteria listed might generate an understanding of what the basic methodological standard is a qualitative study should be able to reach. However, a study may still be judged to have followed the appropriate procedures for a particular approach, yet may suffer from poor interpretation and offer little insight into the phenomenon at hand. Consequently, another study may be flawed in terms of transparency of methodological procedures and yet offer a compelling, vivid and insightful narrative, grounded in the data (Dixon-Woods et al, 2004). Defining fatal flaws and balancing assessment against the weight of a message remains a difficult exercise in the assessment of qualitative studies. As in quantitative research, fatal flaws may depend on the specific design or method chosen (Booth, 2001). This issue needs further research.

Section 2: Stages in the appraisal of qualitative research

Debates in the field of quality assessment of qualitative research designs are centred around a more theoretical approach to evaluating the quality of studies versus an evaluation of the technical adequacy of a research design. How far criteria-based, technical approaches offer significant advantages over expert intuitive judgement in assessing the quality of qualitative research is being challenged by recent evidence indicating that checklist-style approaches may be no better at promoting agreement between reviewers (Dixon-Woods, 2007). However, these appraisal instruments might succeed better in giving a clear explanation as to why certain papers have been excluded. Given the fact that few studies are completely free from methodological flaws, both approaches can probably complement each other.

Is the use of a critical appraisal instruments sufficient in assessing the quality of qualitative studies enhancing Cochrane intervention reviews?

Three different stages can be identified in a quality assessment exercise: filtering, technical appraisal and theoretical appraisal. The first stage links to the inclusion criteria of study types that should be considered to enhance or extent Cochrane Reviews and requires no specific expertise. The required expertise for the next two stages ranges from a basic understanding of qualitative criteria to be able to critically appraise studies to a more advanced level of theoretical knowledge on certain approaches used.

- **Stage 1: Filtering:**
 Within the specific context of enhancing or extending Cochrane Reviews, and viewing critical appraisal as a technical and paradigmatic exercise, it is worth considering limiting the type of qualitative studies to be included in a systematic review. We suggest restricting included qualitative research reports to empirical studies with a description of the sampling strategy, data collection procedures and the type of data-analysis considered. This should include the methodology chosen and the methods or research techniques opted for, which facilitates the systematic use of critical appraisal as well as a more paradigmatic appraisal process. Descriptive papers, editorials or opinion papers would generally be excluded.
- **Stage 2: Technical appraisal:**
 Critical appraisal instruments should be considered a technical tool to assist in the appraisal of qualitative studies, looking for indications in the methods or discussion section that add to the level of methodological soundness of the study. This judgement determines the extent to which the reviewers may have confidence in the researcher's competence in being able to conduct research that follows established norms (Morse, 2002) and is a minimum requirement for critical assessment of qualitative studies. Criteria include but are not limited to the appropriateness of the research design to meet the aims of the research, rigour of data-collection and analysis, well-conducted and accurate sampling strategy, clear statements of findings, accurate representation of participants' voices, outline of the researchers' potential influences, background, assumptions, justifications of the conclusion or whether or not it flows from the data, value and transferability of the research project etc. For this type of appraisal one needs to have a general understanding of qualitative criteria. Involving a researcher with a qualitative background is generally recommended.
- **Stage 3: Theoretical appraisal:**
 In addition to assessing the fulfillment of technical criteria we suggest a subsequent, paradigmatic approach to judgment, with a focus on the research paradigm used in relation to the findings presented. Although some critical appraisal instruments integrate criteria related to theoretical frameworks or paradigms most of them are pragmatic. These do little to identify the quality of the decisions made, the rationale behind them or the responsiveness or sensibility of the researcher to the data. Therefore, a consideration of other criteria should be considered. This would e.g. include an evaluation of methodological coherence or congruity between paradigms that guide the research project and the methodology and methods chosen, an active analytic stance and theoretical position, investigator responsiveness and openness and verification, which refers to systematically checking and confirming the fit between data gathered and

the conceptual work of analysis and interpretation (Morse et al, 2002). For this type of overall judgment a more in-depth understanding of approaches to qualitative research is necessary. It is therefore recommended that a researcher with experience of qualitative research - who can guide others through the critical appraisal process- is invited. Experienced methodologists may have valuable insights into potential biases that are not at first apparent. It should be mentioned though that the need for a paradigmatic input might depend on the type of synthesis chosen.

The Cochrane Qualitative Research Methods group recommends stage 3 whenever the instrument chosen for stage 2 does not cover for a paradigmatic approach to judgment.

Other considerations include involving people with content expertise for the evaluation exercise. They are believed to give more consistent assessments, which is in line with what the Cochrane Collaboration suggests for the assessment of risk of bias in trials (Oxman et al, 1993).

Section 3: A selection of instruments for quality assessment

A range of appraisal instruments and frameworks is available for use in the assessment of the quality of qualitative research. Some are generic, being applicable to almost all qualitative research designs; others have specifically been developed for use with certain methods or techniques. The instruments also vary with regard to the criteria that they use to guide the critical appraisal process. Some address paradigmatic aspects related to qualitative research, others tend to focus on the quality of reporting more than theoretical underpinnings. Nearly all of them address credibility to some extent. The list with examples presented below is not exclusive with many instruments still in development or yet to be validated and others not yet commonly used in practice. It draws on the findings of a review of published qualitative evidence syntheses (Dixon-Woods et al, 2007) and the ongoing update of it. Reviewers need to decide for themselves which instrument appears to be most appropriate in the context of their review and use this judgement to determine their choice. Researchers with a quantitative background also need to consider an input from a researcher familiar with qualitative research, even when an appraisal instrument suitable for novices in the field is opted for.

Which instruments or frameworks are out there?

- **Checklists embedded in a software program to guide qualitative evidence synthesis:**

Some evidence synthesis organisations have developed and incorporated a checklist in the software they make available to assist reviewers with the synthesis of qualitative findings. Typically, potential reviewers need to register to be able to use it. However, the instruments are also available outside the software program on the websites of both organisations¹.

Examples:

QARI software developed by the Joanna Briggs Institute, Australia

URL: <http://www.joannabriggs.edu.au/services/sumari.php>

Used by: Pearson A, Porritt KA, Doran D, Vincent L, Craig D, Tucker D, Long L, Henstridge V. A comprehensive systematic review of evidence on the structure, process, characteristics and composition of a nursing team that fosters a healthy environment. *International Journal of Evidence-Based Healthcare* 2006; 4(2): 118-59.

Rhodes LG et al. Patient subjective experience and satisfaction during the perioperative period in the day surgery setting: a systematic review. *Int J Nurs Pract* 2006; 12(4): 178-92.

EPPI-reviewer developed by the EPPI Centre, United Kingdom

URL: <http://eppi.ioe.ac.uk/eppireviewer/login.aspx>

Used by: Bradley P, Nordheim L, De La Harpa D, Innvaer S & Thompson C. A systematic review of qualitative literature on educational interventions for evidence-based practice. *Learning in Health & Social Care* 2005; 4(2):89-109.

Harden A, Brunton G, Fletcher A, Oakley A. Teenage pregnancy and social disadvantage: a systematic review integrating trials and qualitative studies. *British Medical Journal* Oct 2009.

- **Other online available appraisal instruments:**

Most of the instruments in this selection are easily accessible and clearly define what is meant by each individual criterion listed. As such, they may be particularly useful if reviewers with little experience of qualitative research are required to complete an assessment.

Examples:

Critical Appraisal Skills Programme (CASP):

<http://www.phru.nhs.uk/Doc Links/Qualitative%20Appraisal%20Tool.pdf>

¹ URL: http://www.joannabriggs.edu.au/cqrmg/tools_3.html. A detailed guide on how to conduct a QARI supported Systematic Review, including a detailed explanation of the 10 critical appraisal criteria, can be found on the JBI-website: http://www.joannabriggs.edu.au/pdf/sumari_user_guide.pdf

URL for instrument on process evaluation:

<http://eppi.ioe.ac.uk/cms/default.aspx?tabid=2370&language=en-US>. Tools exist which help to assess quality along three dimensions: quality of reporting, sufficiency of strategies for increasing methodological rigour, and the extent to which study methods and findings are appropriate to answer the review question (For an example, see Harden et al 2009 study).

Used by: Kane GA et al. Parenting programmes: a systematic review and synthesis of qualitative research. Child Care Health and Development 2007; 33(6): 784-793.

Modified versions of CASP, used by:

Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, Donovan J. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. Social Science and Medicine 2003; 56: 671-84.

Malpass A, Shaw A, Sharp D, Walter F, Feder G, Ridd M, Kessler D. 'Medication career' or "Moral career"? The two sides of managing antidepressants: A meta-ethnography of patients' experience of antidepressants. Soc Sci Med. 2009; 68(1):154-68.

Quality Framework UK Cabinet Office

http://www.gsr.gov.uk/downloads/evaluating_policy/a_quality_framework.pdf

Used by: MacEachen E et al. Systematic review of the qualitative literature on return to work after injury. Scandinavian Journal of Work Environment & Health 2006; 32(4): 257-269.

Evaluation Tool for Qualitative Studies

<http://www.fhsc.salford.ac.uk/hcprdu/tools/qualitative.htm>

Used by: McInnes RJ & Chambers JA. Supporting breastfeeding mothers: qualitative synthesis. Journal of Advanced Nursing 2008; 62(4): 407-427.

- **Checklists developed by academics and commonly used in published qualitative evidence syntheses:** Such checklists have been selected and utilised by other researchers in the specific context of an evidence synthesis.

Examples:

The Blaxter (1996) criteria for the evaluation of qualitative research papers, used by:

Gately C et al. Integration of devices into long-term condition management: a synthesis of qualitative studies. Chronic Illn 2008; 4(2): 135-48.

Khan N et al. Guided self-help in primary care mental health - Meta-synthesis of qualitative studies of patient experience. British Journal of Psychiatry 2007; 191: 206-211.

The Burns' (1989) standard for qualitative research, used by:

Barrosso J, Powell Cope GM. Meta-synthesis of qualitative research on living with HIV infection. Qualitative Health Research 2000; 10: 340-53.

Thorne S, Paterson B. Shifting images of chronic illness. *Image: Journal of Nursing Scholarship* 1998; 30; 173-8.
Hildingh C et al. Women's experiences of recovery after myocardial infarction: a meta-synthesis. *Heart Lung* 2007; 36(6): 410-7.
Howard AF, Balneaves LG, Bottorff JL. Ethnocultural women's experiences of breast cancer: a qualitative meta-study. *Cancer nursing* 2007 30(4): E27-35.

The Popay et al (1998) criteria, used by:

Attree P. Low-income mothers, nutrition and health: a systematic review of qualitative evidence. *Maternal and Child Nutrition* 2005 1(4): 227-240.
Sim J & Madden S. Illness experience in fibromyalgia syndrome: A metasynthesis of qualitative studies." *Social Science & Medicine* 2008; 67(1): 57-67.
Yu D et al. Living with chronic heart failure: a review of qualitative studies of older people. *J Adv Nurs* 2008; 61(5): 474-83.

The Mays & Pope (2000) criteria, used by :

Humphreys A et al. A systematic review and meta-synthesis: evaluating the effectiveness of nurse, midwife/allied health professional consultants. *Journal of Clinical Nursing* 2007; 16(10): 1792-1808.
Metcalf A et al. Family communication between children and their parents about inherited genetic conditions: a meta-synthesis of the research. *Eur J Hum Genet* 2008; 16(10): 1193-200.
Robinson L. & Spilsbury K. Systematic review of the perceptions and experiences of accessing health services by adult victims of domestic violence. *Health Soc Care Community* 2008; 16(1): 16-30.

Section 4: Integrating outcomes of critical appraisal in a systematic review.

In a 'best case' scenario a qualitative synthesis or primary study will achieve a positive assessment or score for each of the criteria against which it has been assessed according to the critical appraisal instrument used. However, this will most likely not be the case for the majority of studies and researchers need to be aware of the fact that the assessment or score might depend on the instrument that has been used, which increases the value of involving a researcher with a qualitative background in the appraisal process. For studies that fail to report sufficient information or it is clear that the study is weak when matched against a certain criterion e.g. because of a methodological flaw a decision needs to be made whether to include the study or not.

How to use and report the critical appraisal outcome?

- To include or exclude a study: In this particular case, only high quality studies are included. The potential risk is that valuable insights are excluded from the synthesis. Studies rated as “low quality” because of methodological flaws or lack of reporting may nevertheless generate new insights, grounded in the data, while methodological sound studies may suffer from poor interpretation of data, leading to an insufficient insight into the phenomenon under study (Dixon-Woods, 2007).² This approach was used by Carlsen et al (2007), who excluded 5 out of 17 studies following quality appraisal.

Potential format of presentation:

Study/Criterion*	Study 1	Study 2	Study 3	Study 4	Study 5
Crit 1	x	/	/	x	/
Crit 2	x	x	/	x	x
Crit 3	x	x	x	x	x
Crit 4	x	?	/	x	x
Crit 5	x	x	x	?	x
Quality rating H/L**	H	Judge!***	L	H	Judge!
Comments		Motivate****	Motivate		Motivate

*Authors may choose to give more weight to certain criteria and use this in their final judgment.

** H/L= High/Low

*** For studies that are clearly on the verge between in- and exclusion researchers a judgement on whether to include or exclude should be made and discussed with potential co-reviewers.

**** Authors should include a motivation for in- or exclusion, particularly for those cases where judgments are being made.

- To give more weight to studies that scored high on quality: In this particular case, all valuable insights remain included. However, it might be complex to report on the findings of the synthesis given the ‘subgroups’ of studies. One strategy to cope with the issue of weighing studies is to report the findings of differently rated studies in separate sections. However, this has an impact on the potential richness of the presented synthesis, especially in those approaches generating new theory based on all of the relevant and illuminating findings. No fixed parameters currently exist to determine the weight of qualitative studies. Reviewers choosing this approach need to evaluate which methodological flaws have a

² The loss of potential valuable studies is less likely if one would opt for a critical appraisal instrument which assesses conceptual depth and breadth of findings as well as methodological rigour. Currently, there is no guidance on how these two aspects might be balanced out. Sensitivity analyses could be considered.

substantial impact on the findings presented. The key issue to consider is the extent to which the quality of reporting and choices made by the authors is acceptable in terms of inclusion of the evidence in the Cochrane review which it aims to enhance.

Potential format of presentation:

Study/ Criterion*	Study 1 Subgroup 1	Study 2 Subgr. 1/2	Study 3 Subgr. 2	Study 4 Subgr. 1	Study 5 Subgr. 1/2
Crit 1	x	/	/	x	/
Crit 2	x	x	/	x	x
Crit 3	x	x	x	x	x
Crit 4	x	?	/	x	x
Crit 5	x	x	x	?	x
Qual. rating H/L	H	Judge!	L	H	Judge!
Comments		Motivate	Motivate		Motivate

Both approaches could benefit from a sensitivity analysis evaluating what happens to the findings of a study when low or high quality studies are removed. Thomas et al (2004) conducted such a sensitivity analysis and found that the findings of three studies rated as low quality did not contradict those from studies of a higher quality. This was confirmed by Noyes' and Popays' (2007) study on directly observed therapy and tuberculosis. The synthesis would have come to the same conclusions with or without their inclusion. It indicates that there might be little to gain from including lower quality studies in a synthesis (Harden, 2008).

- To describe what has been observed without excluding any studies: In this particular case, all potential valuable insights remain included, because the worth of individual studies might only become recognisable at the point of synthesis rather than in the phase of appraisal. In this approach, the responsibility for evaluating the quality of the studies is devolved to the reader from the researcher.

The Cochrane Qualitative Research Methods Group sees value in all of these approaches. However, in line with current Cochrane policy, when conducting a Cochrane Intervention review and integrating qualitative evidence we recommend the two first approaches emphasizing the methodological soundness of studies rather than their contribution to science in general. The decision lies with the review team. Regardless of the approach eventually chosen for the quality assessment stage of the review there is a need to preserve the transparency of the method through careful documentation of decisions made. The convention of using at least two researchers for the quality assessment process is a useful legacy from quantitative-based review processes; not so

much for inter-rater consistency purposes but, at the very least, to open up the data to a broader range of possible interpretations.

Conclusion

Quality assessment of qualitative research studies remains a contested area. While considerable widespread debate continues around the feasibility and utility of critical appraisal it is nevertheless possible to make recommendations within the specific context of informing, enhancing and extending a Cochrane Review. Balancing assessment against the weight of a message is a difficult exercise. In assessing the quality of an original study reviewers should focus on quality of reporting, methodological rigour and conceptual depth and breadth of studies. Review Teams should deliberately seek indications that demonstrate rigour truth- and trustworthiness in the studies (**section 1**). Filtering, technical appraisal and theoretical appraisal are the three main stages in a critical appraisal exercise. Review teams need to make sure that they have the necessary expertise on board to be able to complete each phase of the critical appraisal exercise (**section 2**). In choosing an assessment instrument review teams need to consider the appropriateness of their choice in the context of their review and be aware of the fact that whether or not a study meets the standard might depend on the instrument used (**section 3**). Review teams can opt to use the outcome of their critical appraisal to in- or exclude studies, to give high quality studies more weight or to inform themselves without excluding any. Regardless of the approach eventually chosen there is a need to preserve the transparency of the method through careful documentation of decisions made. The use of at least two researchers to complete the critical appraisal process is recommended (**section 4**).

References

Blaxter M. Criteria for evaluation of qualitative research. *Medical Sociology News* 1996; 22: 68-71.

Booth A. Cochrane or cock-eyed? How should we conduct systematic reviews of qualitative research? Qual EBP conference, Coventry university, may 14-16, 2001.

Burns N. Standards for Qualitative Research. *Nurs Sci Q* 1989; 2(44): 44-52.

Carlsen B, Glenton C, Pope C. Thou shalt versus thou shalt not: a meta-synthesis of GPs' attitudes to clinical practice guidelines. *Br J Gen.Pract* 2007;57:971-8.

Cohen DJ & Crabtree BF. Evaluative criteria for qualitative research in health care; controversies and recommendations. *Annals of Fam. Med* 2008; 6 (4): 331-39.

Dixon-Woods M, Booth A, Sutton AJ. Synthesizing qualitative research: a review of published reports. *Qual Res* 2007; 7: 375-421.

Dixon-Woods M, Shaw RL, Agarwal S, Smith JA. The problem of appraising qualitative research. *QSHC* 2004;13: 223-225.

Dixon-Woods M, Booth A, Sutton AJ. Synthesizing qualitative research: a review of published reports. *Qualitative Research* 2007; 7(3): 375-422.

Dixon-Woods M, Sutton AF, Shaw RL, Miller T, Smith J, Young B, Bonas S, Booth A, Jones DR. Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. *Journal of Health Service Research and Policy* 2007; 12: 42-7.

Edwards A, Elwyn G, Hood K, Rollnick S. Judging the 'weight of evidence' in systematic reviews: introducing rigour into the qualitative overview stage by assessing Signal and Noise. *J Eval Clin Pract.* 2000 May;6(2):177-84.

Lincoln YS, Guba EG. *Naturalistic Inquiry*, Sage, Newbury Park, CA, 1985 .

Guba EG. & Lincoln YS. *Fourth Generation evaluation*. Newbury Park. CA: Sage. 1989.

Harden A. *Critical Appraisal and Qualitative Research: Exploring sensitivity analysis*. ESRC Research Methods Festival, St. Catherine's College Oxford, 30th June-3rd July, 2008.

Hill A & Spittlehouse C. What is critical appraisal? *Evidence-Based Medicine* 2003; 3 (2): 1-8.

Noyes J & Popay J. Directly observed therapy and tuberculosis: how can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis. *Journal of Advanced Nursing* 2007; 57(3), 227–243.

Morse JM, Barrett M, Mayan M, Olson K, Spiers J. Verification strategies for establishing reliability and validity in qualitative research. *Int J of Qual Meth* 2002; 1 (2):spring.

Mays N, Pope C. Qualitative research in health care Assessing quality in qualitative research. *BMJ* 2000;320:50-52.

Noyes J, Popay J, Pearson A, Hannes K, Booth A. Chapter 20: Qualitative research and Cochrane reviews. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane-handbook.org.

Oxman AD, Guyatt GH. The science of reviewing research. *Annals of the New York Academy of Sciences* 1993; 703: 125-133.

Popay, J., & Williams, G. (1998). Qualitative research and evidence based healthcare. *Journal of Research in Social Medicine*, 91(Suppl 35), 32–37.

Popay J, Rogers A, Williams G. Rationale and Standards for the Systematic Review of Qualitative Literature in Health Care. *Qual Health Research* 1998; 8: 341-51.

Spencer L, Ritchie J, Lewis J et al. et al. *Quality in Qualitative Evaluation: a framework for assessing research evidence: a framework for assessing research evidence*. Government chief social researcher's office 2003.

Thomas J, Harden A, Oakley Am Oliver S, Sutcliffe K, Rees R, Brunton G, Kavanagh J. Integrating qualitative research with trials in systematic reviews. *BMJ* 2004; 328: 1010-1012.