

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7523973>

# Appraising the quality of research

Article *in* Midwifery · July 2006

DOI: 10.1016/j.midw.2005.05.004 · Source: PubMed

---

CITATIONS

137

---

READS

1,393

2 authors:



[Denis Walsh](#)

University of Nottingham

25 PUBLICATIONS 1,292 CITATIONS

[SEE PROFILE](#)



[Soo Downe](#)

University of Central Lancashire

186 PUBLICATIONS 2,060 CITATIONS

[SEE PROFILE](#)



# Appraising the quality of qualitative research

Denis Walsh, MA RM, DPSM (Senior Lecturer)<sup>a,\*</sup>,  
Soo Downe, RM, BSc, PhD (Professor of Midwifery Studies)<sup>b</sup>

<sup>a</sup>University of Central Lancashire, 366 Hinckley Road, Leicester LE3 OTN, UK

<sup>b</sup>University of Central Lancashire, Preston, UK

Received 18 November 2002; received in revised form 8 August 2004 and 9 May 2005; accepted 13 May 2005

## KEYWORDS

Meta-synthesis;  
Quality;  
Qualitative research;  
Checklist

**Summary** In the process of undertaking a meta-synthesis of qualitative studies of free-standing midwife-led units, the authors of this paper encountered a number of methodologically and epistemologically unresolved issues. One of these related to the assessment of the quality of qualitative research. In an iterative approach to scoping this issue, we identified eight existing checklists and summary frameworks. Some of these publications were opinion based, and some involved a synthesis of pre-existing frameworks. None of them provide a clear map of the criteria used in all their reviewed papers, and of the commonalities and differences between them. We critically review these frameworks and conclude that, although they are epistemologically and theoretically dense, they are excessively detailed for most uses. In order to reach a workable solution to the problem of the quality assessment of qualitative research, the findings from these frameworks and checklists were mapped together. Using a technique we have termed a 'redundancy approach' to eliminate non-essential criteria, we developed our own summary framework. The final synthesis was achieved through reflexive debate and discussion. Aspects of this discussion are detailed here. The synthesis is clearly rooted in a subjectivist epistemology, which views knowledge as constructed and hermeneutic in intent, encompassing individual, cultural and structural representations of reality.

© 2005 Elsevier Ltd. All rights reserved.

## Background

Narratives of womens' experiences of midwifery care have been published since at least the 1960s (Kitzinger, 1962); however, the first explicitly

research-based account of the nature of English midwifery practice was not completed until 1983 (Kirkham, 1983). Over the past 2 decades, qualitative research in maternity care has gained increasing exposure and credibility. This reflects a growing interest in this paradigm in the health services, as it sheds light on the environment and culture of care (Hunt and Symonds, 1995), on why clinicians

\*Corresponding author.

E-mail address: Denis.walsh@ntlworld.com (D. Walsh).

practice as they do (Grol, 1997) and on the client's experience (Liamputtong and Naksook, 2003).

The increasing volume of qualitative papers has prompted attempts at meta-synthesis. This technique has similarities with, as well as important differences from, meta-analysis of quantitative studies. There have been at least four meta-synthesis papers published about aspects of the maternity experience: Beck's (2002a) examination of mothering multiples (twins, triplets, quadruplets), and her (Beck, 2002b) account of postpartum depression; Clemmens' (2003) exploration of adolescent motherhood; and Kennedy et al.'s (2003) review of midwifery practice in the USA.

The technique of meta-synthesis is still developing and contested areas remain. The authors of this paper encountered some of these in the process of undertaking a meta-synthesis of qualitative studies of free-standing midwife-led units (MLUs), and have written about them elsewhere (Walsh and Downe, 2005). In particular, earlier examples of meta-synthesis were equivocal in the extent to which individual studies should be thoroughly appraised for quality before inclusion. Britten et al. (2002) took a pragmatic approach. They prioritised the need for a worked example of the stages of meta-synthesis over a rigorous selection of studies on the basis of pre-stated quality criteria. Although we do not dispute the value of this case-study approach, it does raise for us the possibility that meta-synthesis of methodologically flawed studies may result in flawed meta-synthesis. In keeping with our particular stance, we decided to rigorously appraise studies first before submitting them to the meta-synthesis technique. This required agreement on criteria to judge rigor. It is the process we underwent in establishing these criteria that is detailed in this paper.

We are aware that our position on quality inclusion is based on a philosophical assumption that qualitative research can be flawed, and that this assumption would be vigorously resisted by some researchers working in this area. Debates about appropriate criteria for appraising qualitative research have existed for at least the last 15 years (Holloway and Wheeler, 1996; Perakyla, 1997). There are a number of reasons for this continuing lack of agreement. The first, and most fundamental, has to do with the nature of evidence or knowledge produced by this method.

Qualitative research has strong links with interpretivism. The epistemological stance arising from this sees knowledge as a socially produced construct (Crotty, 1998). This contrasts with positivism, an earlier and more dominant epistemology that understands knowledge as accessible through

what can be observed and what can be rationally deduced through reason. The positivist epistemology believes knowledge to be objective and 'true', in contrast to interpretivists' points of view, which understand knowledge to be constructed, and therefore not as stable or as objective. In between these positions are the post or neo positivists whose critique embraces both the decontextualisation of positivist research and the political inertia of relativist epistemologies (Harding, 1991; Parsons, 1994). It is fairly easy to see how a positivist stance can rather straightforwardly evolve techniques to establish its knowledge claims. The development of the technique of the randomised controlled trial is a prime illustration of this point.

The limitations of this kind of knowledge as the sole means of addressing the complexity of health care are becoming apparent. Clinicians and managers are increasingly keen to know how qualitative insights can be used in health care. As a consequence, they require guidance to help them to discriminate between 'good' and 'bad' studies. Many clinicians have come to the application of research appraisal techniques through concepts such as validity and reliability. Consequently, they seek to apply the same criteria to qualitative work (Tobin and Begley, 2004). However, the different epistemological status of most qualitative research makes the indiscriminate transferral of these criteria inappropriate. There is, however, no clear agreement on an alternative more appropriate stance.

Qualitative research covers a broad church of philosophical positions. These range from extreme post-modernist relativism, in which there are a multiplicity of personal truths that change as we change (Fox, 1993), to a critical realist view that sees underlying processes, such as economic or political structures, as causing real effects, such as oppression and disempowerment (Harvey, 1990). In between, there are a variety of other positions that variously emphasise the role of language (symbolic interactionists), interpretation and meanings (hermeneutics), lived experience (phenomenology) and group behaviours and beliefs (ethnography), to name just a few. It is, therefore, not surprising that consensus is hard to reach on the topic of appraisal. Indeed, some qualitative writers argue that they are simply telling their particular story and it is up to the reader to assign relevance. This somewhat extreme position disallows any method of establishing 'rightness' or legitimacy. The intent in this case is to reject any attempt to convince or compel others into accepting one version of events as authoritative, in distinct contrast to the positivist intent (Rolfe, 2000).

However, this fluid approach is unacceptable to those who are seeking direction for health-care interventions, or who desire to improve the quality of care. As [Murphy et al. \(1998\)](#) have stated:

Some argue that...the very idea of criteria is incompatible with... (the... anti-realist assumptions... (of qualitative research)... We suggest that this position is unnecessarily constraining ...if the findings of research cannot be taken to represent even an approximation of the truth... why should commissioners... fund... such research (p. 10).

In an attempt to square the apparent epistemological circle arising from these divergent philosophical stances, we undertook a scoping review of current frameworks for the assessment of the quality of qualitative research. We then appraised and synthesised the resulting frameworks to make a checklist that was useable in practice, as well as being adequately comprehensive.

## Scoping of the issue

### Literature review

In the process of scoping this issue, we came across a number of checklists for appraising qualitative research. This occurred through an iterative process akin to [Bates' \(1989\)](#) 'berrypicking' model, rather than through a systematic search of the literature. This approach reflects 'real world' search patterns, where the retrieval of one paper leads to others. Four different checklists were found in journals ([Popay and Rogers, 1998](#); [Mays and Pope, 2000](#); [Yardley, 2000](#); [Cesario et al., 2002](#)). Networking via an email group led us to two further sources, both of which were systematic approaches commissioned to synthesise a definitive framework ([Murphy et al., 1998](#); [Spencer et al., 2003](#)). Another source was discovered via the UK Critical Appraisal Skills Programme (CASP), which trains service users and clinicians together in research report appraisal ([CASP, 1999](#)). Finally, a search of four databases discovered one more synthesised framework ([Sandelowski and Barroso, 2002](#)). In total, eight checklists were identified through this method (see [Table 1](#) for features of included studies). Some of these publications were opinion based ([Yardley, 2000](#)), and some involved a synthesis of pre-existing frameworks ([Murphy et al., 1998](#)). Three summary frameworks were located ([Murphy et al., 1998](#); [Sandelowski and Barroso, 2002](#); [Spencer et al., 2003](#)). They intro-

duced us to a further extensive list of existing checklists, indicating an apparent need in this area, and a signal failure to reach any form of synthesis. To illustrate this point, [Spencer et al. \(2003\)](#) located 29 different checklists in their search.

Two of the summary lists provide narrative syntheses, tracking the process of how criteria were eventually distilled ([Sandelowski and Barroso, 2002](#); [Spencer et al., 2003](#)). However, none of them provide a clear map of the criteria used in all their reviewed papers, and of the commonalities and differences between them. All three summary frameworks are dense and lengthy, and unlikely, in our view, to be widely adopted. It is of interest to locate their origins and explore the context within which they were developed.

[Murphy et al. \(1998\)](#) were commissioned by the Health Technology Assessment Board (HTAB) to review qualitative research generally, and to assess its value and relevance in technology assessment. The HTAB is a UK-based group that usually operates to strict quantitative criteria in appraising new technologies. Part of the work included reviewing current criteria for appraising the quality of qualitative research. Owing to the complexity of the undertaking, the authors produced a number of guiding principles, rather than a checklist.

The Cabinet Office of the UK government commissioned [Spencer et al. \(2003\)](#) to establish clear and unambiguous criteria that could be used across the whole spectrum of health and social care. The exercise was detailed, comprehensive and exhaustive. Adapting a modified Delphi technique, they consulted experts from various fields and honed down criteria over a 12-month period. They produced an extensive table, with a level of detail that seemed to reflect the inability of the experts to reconcile their differing emphases. In our view, their final table is unlikely to be widely used because it is unwieldy and cumbersome.

Finally, [Sandelowski and Barroso \(2002\)](#) undertook their work for similar reasons to us. They were conducting a meta-synthesis of research into women with HIV/AIDS, and had gathered a team of qualitative researchers to appraise the studies in this area. Differences over criteria and how they should be applied were immediately manifested. In response, they engaged the experts in a Delphi-style exercise to list relevant criteria, and to weight them regarding level of importance for assessing the integrity of qualitative papers. Once again, their appraisal template is lengthy. [Sandelowski and Barroso \(2002\)](#) do make the interesting point that failure of one criterion should not negate the value of the paper. In other words, they counsel the reflexive use of the criteria so that important

Table 1 Features of included studies.

| Authors, date and Country          | Professional field | Orientation  | Method of derivation                   | Number of studies ( <i>n</i> ) included | Method of testing  | Findings and conclusions                    | Comments   |
|------------------------------------|--------------------|--------------|--|---|--|---|--|
| Mayes and Pope, 1995, UK           | Healthcare         | Interpretive | Opinion                                | 1                                       | Not done   | Questions to guide process                  | Aimed at medical staff; uncomplicated; emphasis on utility and overview              |
| Murphy et al., 1998, UK            | Healthcare         | Interpretive | Systematic review                      | Multiple <i>n</i> not stated            | Not done   | Comprehensive list of areas to be addressed | Thorough review of literature  |
| CASP, 1999, UK                     | Healthcare         | Not stated   | Opinion                                | 1                                       | No   | Checklist                                   | List-based without any engagement with interpretivist paradigm                       |
| Popay and Rogers 1998, UK          | Healthcare         | Interpretive | Opinion                                | 1                                       | Not done   | Questions to guide process                  | Helpful overview that engages with interpretivist paradigm                           |
| Yardley, 2000, UK                  | Healthcare         | Interpretive | Opinion                                | 1                                       | Not done   | Minimalist table                            | Authors reluctant to provide definitive guidance because of epistemological concerns |
| Cesario et al., 2002, USA          | Healthcare         | Positivist   | Opinion                                | 1                                       | Not done   | Prescriptive table                          | Not convincing as positivist bias clear  |
| Sandelowski and Barroso, 2002, USA | Healthcare         | Interpretive | Literature review and expert consensus | Multiple <i>n</i> not stated            | Yes: iteratively developed through expert panel review, testing on located studies, and retesting by the experts on a further set of theoretically sampled studies | Lengthy template                            | Authors detail clearly how they arrived at template                                  |
| Spencer et al., 2003, UK           | Social care        | Interpretive | Systematic review and Expert consensus | 29                                      | Yes, iteratively developed through stakeholder interviews and workshops, and then tested on eight theoretically sampled studies                                    | Lengthy template                            | Authors detail clearly how they arrived at template                                  |

findings are not compromised by apparent lack of rigor in another area, which is not seminal to the overall integrity of the paper.

The three existing summary reports detailed above do seem to be the most comprehensive work to date in this area. None of them, however, provided a simple summary template for assessing the quality of qualitative research. We decided that our original sample of eight lists, including the three summary papers, provided sufficient data to make an attempt at creating a workable and comprehensive guide. The next section describes the process we undertook.

### Mapping exercise

We initially tabulated the characteristics of each of the studies in our review. We then mapped together

the characteristics given in all the included papers (Table 2), sorting them by the number of checklists in which they appeared.

### Synthesising accounts

Both authors independently attempted a synthesis from Table 2 before coming together to discuss our decisions. We undertook this by looking for redundancy in the included criteria. We examined each included item to see if its exclusion would change our overall judgement on the meaningfulness and applicability of a piece of qualitative research. This process is made more transparent in Table 3, in which criteria are categorised as 'essential', 'desirable' or 'optional'. The table represents something of an audit trail of how decisions were arrived at. If we both agreed that

**Table 2** Criteria by checklists.

|                                 | Criteria  | n | A  | B  | C  | D  | E  | F  | G  | H  |
|---------------------------------|---|---|----|----|----|----|----|----|----|----|
| Common to all                   | Method/design apparent and appropriate  | 8 | *  | *  | *  | *  | *  | *  | *  | *  |
|                                 | Data collection methods apparent and appropriate                                | 8 | *  | *  | *  | *  | *  | *  | *  | *  |
|                                 | Analytic approach apparent and appropriate                                      | 8 | *  | *  | *  | *  | *  | *  | *  | *  |
|                                 | Context described sufficiently  | 8 | *  | *  | *  | *  | *  | *  | *  | *  |
|                                 | Relevance and transferability evident   | 8 | *  | *  | *  | *  | *  | *  | *  | *  |
| Found in at least 50% of papers | Purpose/aim/problem stated  | 5 | *  | *  | *  | *  |    | *  |    |    |
|                                 | Theoretical/epistemological underpinning evident                                | 7 | *  | *  | *  |    | *  | *  | *  | *  |
|                                 | Sampling strategy appropriate   | 6 | *  | *  | *  | *  |    | *  |    | *  |
|                                 | Use of triangulation discussed  | 6 | *  | *  | *  | *  | *  |    | *  |    |
|                                 | Researcher reflexivity demonstrated   | 7 | *  | *  | *  | *  | *  |    | *  | *  |
|                                 | 'Subjective' meanings/ phenomena treated as data                                | 5 | *  | *  |    | *  |    |    | *  | *  |
|                                 | Attention given to negative/dissonant cases                                     | 5 | *  | *  | *  |    | *  |    | *  |    |
|                                 | Member checking done  | 7 | *  | *  | *  | *  | *  | *  |    | *  |
|                                 | Analysis repeated with another researcher                                       | 4 | *  | *  |    | *  | *  |    |    |    |
|                                 | Evidence data saturation reached  | 5 | *  |    |    |    | *  | *  | *  | *  |
|                                 | Use of data to support interpretation   | 5 | *  | *  | *  |    |    | *  | *  | *  |
|                                 | Ethical dimensions discussed  | 5 | *  | *  | *  | *  |    | *  | *  | *  |
|                                 | Audit trail apparent  | 5 | *  | *  | *  |    | *  | *  |    | *  |
|                                 | Discussion of how findings add to existing knowledge                            | 6 | *  | *  | *  |    | *  | *  |    | *  |
|                                 | Researcher/participant relationship/partnership/fair dealing attended to        | 5 |    | *  | *  | *  | *  |    | *  |    |
| Found in few papers             | Evidence of thorough relevant literature search (at any stage of the report)    | 3 | *  | *  |    |    |    |    | *  |    |
|                                 | Written record clear and logical  | 2 | *  | *  |    |    |    |    |    |    |
|                                 | Time frame mentioned  | 2 | *  | *  |    |    |    |    |    |    |
|                                 | Problems encountered/limitations of study discussed                             | 2 | *  | *  |    |    |    |    |    |    |
| Found in one paper              | Evidence of adaptation of design in response to changes in setting during study | 1 |    |    |    |    |    |    |    | *  |
|                                 | Researcher suitability for undertaking study                                    | 1 | *  |    |    |    |    |    |    |    |
|                                 | 'Hawthorne' effects discussed   | 1 |    |    |    |    |    | *  |    |    |
|                                 | Suggests further directions for research  | 1 |    |    |    | *  |    |    |    |    |
|                                 | Number of elements identified   |   | 25 | 25 | 17 | 16 | 16 | 15 | 15 | 14 |

A, Spencer et al., 2003; B, Sandelowski and Barroso, 2002; C, Mays and Pope, 1995; D, CASP, 1999; E, Murphy et al., 1999; F, Cesario et al., 2002; G, Yardley, 2000; H, Popay and Rogers, 1998; n, numbers of checklists stating this criterion.

the exclusion would not change the final judgment, it was left out. The final act was to draw up a definitive checklist, structured into three columns, namely stages, essential criteria and specific prompts (Table 4). Although some of the resulting criteria may be self-evident (i.e. sample and setting, and data collection methods), others may seem less obviously fundamental. We discuss three of these less obvious aspects below.

### Identification of method that is consistent with research intent

Some of the included checklists did not require researchers to specify the method they used, being content with a catchall phrase like 'descriptive' or 'interpretive' to position the approach within the qualitative arena (CASP, 1999; Yardley, 2000). In our

view, this is inadequate, because specific methods have evolved with different emphases that are particularly suited to particular spheres of investigation. If the culture of an environment is being explored, then ethnography is most appropriate as method. If the focus is on an in-depth exploration of subjective experience, then phenomenology would be suitable. If 'talk' or dialogue is under scrutiny, then discourse analysis is indicated. Where the nature of the particular method used is not recognised by the researchers, there is a risk of a certain fuzziness that may extend to data collection methods and analysis. The theoretical constructs underpinning each specific method entail a certain set of data collection methods and analytic approaches. Although these may overlap between methodological approaches, a lack of clarity at the outset may lead to inappropriate choices and conclusions later on in the work.

**Table 3** First stage of mapping exercise.

|                                 | Criteria  | Essential | Desirable | Optional |
|---------------------------------|---|-----------|-----------|----------|
| Common to all                   | Method/design apparent and appropriate  | *         |           |          |
|                                 | Data collection methods apparent and appropriate                                | *         |           |          |
|                                 | Analytic approach apparent and appropriate                                      | *         |           |          |
|                                 | Context described sufficiently  | *         |           |          |
|                                 | Relevance and transferability evident   | *         |           |          |
| Found in at least 50% of papers | Purpose/aim/problem stated  | *         |           |          |
|                                 | Theoretical/epistemological underpinning evident                                |           | *         |          |
|                                 | Sampling strategy appropriate   | *         |           |          |
|                                 | Use of triangulation discussed  |           |           | *        |
|                                 | Researcher reflexivity demonstrated   | *         |           |          |
|                                 | 'Subjective' meanings/ phenomena treated as data                                | *         |           |          |
|                                 | Attention given to negative/dissonant cases                                     |           | *         |          |
|                                 | Member checking done  | *         | *         |          |
|                                 | Analysis repeated with another researcher                                       |           | *         |          |
|                                 | Evidence data saturation reached  |           | *         |          |
|                                 | Use of data to support interpretation   |           |           |          |
|                                 | Ethical dimensions discussed  |           |           |          |
|                                 | Audit trail apparent  | *         |           |          |
|                                 | Discussion of how findings add to existing knowledge                            | *         | *         |          |
|                                 | Researcher/participant relationship/partnership/fair dealing attended to        |           | *         |          |
| Found in a few papers           | Evidence of thorough relevant literature search (at any stage of the report)    |           |           |          |
|                                 | Written record clear and logical  |           | *         |          |
|                                 | Time frame mentioned  | *         |           | *        |
|                                 | Problems encountered/limitations of study discussed                             |           |           |          |
| Found in one paper              | Evidence of adaptation of design in response to changes in setting during study |           |           | *        |
|                                 | Researcher suitability for undertaking study                                    | *         |           | *        |
|                                 | 'Hawthorne' effects discussed   |           |           | *        |
|                                 | Suggests further directions for research  |           | *         |          |

**Table 4** Summary criteria for appraising qualitative research studies.

| Stages            | Essential criteria   | Specific prompts  |
|-------------------|--|---|
| Scope and purpose | Clear statement of, and rationale for, research question/aims/purposes | <ul style="list-style-type: none"> <li>• Clarity of focus demonstrated</li> <li>• Explicit purpose given, such as descriptive/explanatory intent, theory building, hypothesis testing</li> <li>• Link between research and existing knowledge demonstrated</li> </ul>   |
|                   | Study thoroughly contextualised by existing literature                 | <ul style="list-style-type: none"> <li>• Evidence of systematic approach to literature review, location of literature to contextualise the findings, or both</li> </ul>   |
| Design            | Method/design apparent, and consistent with research intent            | <ul style="list-style-type: none"> <li>• Rationale given for use of qualitative design</li> <li>• Discussion of epistemological/ontological grounding</li> <li>• Rationale explored for specific qualitative method (e.g. ethnography, grounded theory, phenomenology)</li> <li>• Discussion of why particular method chosen is most appropriate/sensitive/relevant for research question/aims</li> <li>• Setting appropriate</li> </ul>  |
|                   | Data collection strategy apparent and appropriate                      | <ul style="list-style-type: none"> <li>• Were data collection methods appropriate for type of data required and for specific qualitative method?</li> <li>• Were they likely to capture the complexity/diversity of experience and illuminate context in sufficient detail?</li> <li>• Was triangulation of data sources used if appropriate?</li> </ul>  |
| Sampling strategy | Sample and sampling method appropriate                                 | <ul style="list-style-type: none"> <li>• Selection criteria detailed, and description of how sampling was undertaken</li> <li>• Justification for sampling strategy given</li> <li>• Thickness of description likely to be achieved from sampling</li> <li>• Any disparity between planned and actual sample explained</li> </ul>   |
| Analysis          | Analytic approach appropriate  | <ul style="list-style-type: none"> <li>• Approach made explicit (e.g. Thematic distillation, constant comparative method, grounded theory)</li> <li>• Was it appropriate for the qualitative method chosen?</li> <li>• Was data managed by software package or by hand and why?</li> <li>• Discussion of how coding systems/conceptual frameworks evolved</li> <li>• How was context of data retained during analysis</li> <li>• Evidence that the subjective meanings of participants were portrayed</li> <li>• Evidence of more than one researcher involved in stages if appropriate to epistemological/theoretical stance</li> <li>• Did research participants have any involvement in analysis (e.g. member checking)</li> <li>• Evidence provided that data reached saturation or discussion/rationale if it did not</li> <li>• Evidence that deviant data was sought, or discussion/rationale if it was not</li> </ul> |
| Interpretation    | Context described and taken account of in interpretation               | <ul style="list-style-type: none"> <li>• Description of social/physical and interpersonal contexts of data collection</li> <li>• Evidence that researcher spent time 'dwelling with the data', interrogating it for competing/alternative explanations of phenomena</li> </ul>  |
|                   | Clear audit trail given  | <ul style="list-style-type: none"> <li>• Sufficient discussion of research processes such that others can follow 'decision trail'</li> </ul>  |

Table 4 (continued)

| Stages                        | Essential criteria                               | Specific prompts  |
|-------------------------------|--|---|
|                               | Data used to support interpretation              | <ul style="list-style-type: none"> <li>• Extensive use of field notes entries/verbatim interview quotes in discussion of findings</li> <li>• Clear exposition of how interpretation led to conclusions</li> </ul>   |
| Reflexivity                   | Researcher reflexivity demonstrated              | <ul style="list-style-type: none"> <li>• Discussion of relationship between researcher and participants during fieldwork</li> <li>• Demonstration of researcher's influence on stages of research process</li> <li>• Evidence of self-awareness/insight</li> <li>• Documentation of effects of the research on researcher</li> <li>• Evidence of how problems/complications met were dealt with</li> </ul>  |
| Ethical dimensions            | Demonstration of sensitivity to ethical concerns | <ul style="list-style-type: none"> <li>• Ethical committee approval granted</li> <li>• Clear commitment to integrity, honesty, transparency, equality and mutual respect in relationships with participants</li> <li>• Evidence of fair dealing with all research participants</li> <li>• Recording of dilemmas met and how resolved in relation to ethical issues</li> <li>• Documentation of how autonomy, consent, confidentiality, anonymity were managed</li> </ul>  |
| Relevance and transferability | Relevance and transferability evident            | <ul style="list-style-type: none"> <li>• Sufficient evidence for typicality specificity to be assessed</li> <li>• Analysis interwoven with existing theories and other relevant explanatory literature drawn from similar settings and studies</li> <li>• Discussion of how explanatory propositions/emergent theory may fit other contexts</li> <li>• Limitations/weaknesses of study clearly outlined</li> <li>• Clearly resonates with other knowledge and experience</li> <li>• Results/conclusions obviously supported by evidence</li> <li>• Interpretation plausible and 'makes sense'</li> <li>• Provides new insights and increases understanding</li> <li>• Significance for current policy and practice outlined</li> <li>• Assessment of value/empowerment for participants</li> <li>• Outlines further directions for investigation</li> <li>• Comment on whether aims/purposes of research were achieved</li> </ul> |

### Researcher reflexivity

Over the past 20 years, researcher reflexivity has become increasingly significant for qualitative researchers. Clifford and Marcus (1986) made a seminal contribution to the reflexivity 'turn' in highlighting the role of the researcher in constructing the written account with their expositions on recording ethnographic accounts of culture. Around the same time, Strathern's (1988) highlighted the Eurocentric bias in anthropological studies of

Melanesian cultures, highlighting gender in particular as a fluid code. These authors brought to centre stage the need to be reflexive about an investigator's presuppositions. We agreed strongly that researcher reflexivity is a key tenet of qualitative research, lending it an authenticity and honesty that is distinctive. It is an area in which the divide between qualitative and quantitative research is most obvious, the latter either ignoring the effect of the researcher on the researched, or repackaging it as value-free

neutrality (Christians, 2000). By way of contrast, even in early anthropological ethnographies, the origins of reflexivity can be gleaned. Malinowski recognised the importance of representing indigenous cultures through their own eyes (Kuper, 1983), prefacing the later 'etic' (outsider)/ 'emic' (insider) dynamic that has richly fed into both analysis and results of ethnographies since Hammersley and Atkinson (1995).

Sadly, researcher reflexivity is often culled from journal publications that restrict wordage from original papers. We argue that it is imperative to publish some reflexive content so that the reader can sense how the researcher shaped the entire project, and, in particular, the interpretation of findings. Even if it is not available in published accounts, it is our view that those undertaking meta-synthesis should be able to obtain this aspect of the work from the researchers, as it is paramount for judging the integrity of the work. Given this view, we agreed that the criterion of researcher reflexivity was fundamental to our checklist.

### Ethical dimensions

Many qualitative researchers, and especially those working in critical approaches, such as feminism and disability research, have championed concern with equality in the researcher/participant relationship. Oakley (1993) first questioned whether research was 'on' or 'with' women in challenging the traditional hierarchical relationship of the researcher and the researched. Sensitivity to, and respect for, the status and integrity of research subjects is most visible now through mandatory ethical approval procedures. These are based on the protection of individuals from harm through guarantees of confidentiality, anonymity and informed written consent. Qualitative researchers often take this further than the 'one-off' approval process at the beginning of studies by explicitly keeping participants informed at all stages of the research process, and by attempts to ensure that participants encounter respect, transparency and openness (Birch et al., 2002) This emphasises an ethical underpinning to all research endeavour, beyond mere adherence to ethical procedures. However, anthropology and sociology have also endorsed techniques such as covert research. Work that is undertaken on the boundaries of ethical acceptability is most in need of ethical transparency and justification in its reporting. Although some of the checklists and frameworks we examined clearly imply that ethical considerations are

imperative (Popay and Rogers, 1998; Murphy et al., 1998), others explicitly state this (CASP, 1999; Yardley, 2000; Cesario et al., 2002; Sandelowski and Barroso, 2002; Spencer et al., 2003). We believe that it should be an explicit component of good-quality qualitative research.

Although consensus was straightforward for most elements in the checklist, some areas caused considerable reflection and debate. The issues we raised between us are summarised in the next section of this paper.

### Literature reviews

After some debate, we reached the conclusion that a well-conducted literature review was essential, good quality qualitative research. We acknowledge that tension exists regarding the extent to which a researcher should search out and establish the state of knowledge about the topic being explored before undertaking the primary data collection and analysis. Indeed, grounded theorists caution against this approach. Glaser and Strauss (1967) argue that it will pre-empt and unduly influence emerging theory, and therefore stifle original insights into the area. Exposing existing knowledge may in fact delimit alternative explanations of phenomena. However, because researchers approach an area of enquiry with existing presuppositions based on their personal history and interest in the topic, such detachment is at best unrealistic and at worst dishonest. The importance of a reflexive disposition renders a position of 'knowing ignorance' untenable in our view.

The question then becomes how systematic does a literature review have to be? Limitations are often imposed by the accessibility of qualitative sources. Unlike much published quantitative research, qualitative papers appear more frequently in books and book chapters, both of which are under-represented in databases. In addition, commonly accessed medical databases, such as Medline, have until recently had limited links to qualitative research papers. As an example, Medline did not index the seminal journal, *Qualitative Health Research*, until 1999 (Paterson et al., 2001), and the word 'qualitative' has only recently been added as a primary search term (Barroso et al., 2003).

We acknowledge that a literature review may be iterative in the context of the emergent nature of qualitative research. In addition, journal wordage requirements may result in the inclusion of only minimal levels of detail around aspects of the work, such as the literature review. However, we feel

strongly that a full account of a qualitative study needs to be transparently comprehensive in its incorporation of existing literature, whether this is before the onset of the study or to contextualise the findings.

### Analytic approach

The analytic approach in qualitative research has been the subject of much debate ([Miles and Huberman, 1994](#); [Morse, 1994](#); [Sandelowski et al., 1997](#)). Some authors critique a lack of explicitness about how themes are distilled and how theories emerge ([Murphy et al., 1998](#)). It is this apparent 'leap of faith' that grounded theory seeks to audit. We debated the value of specific aspects of a trustworthy analysis. This debate included discussion on the value, or otherwise, of the use of qualitative software packages in providing a comprehensive audit trail. Whatever the chosen approach to synthesis, it may be difficult to capture the inductive steps diagrammatically, and integrity of findings has to be argued in the write-up. Acknowledging dissonant or deviant data may have a key part to play here, as the explicit recognition that it exists indicates a rigorous and transparent approach to analysis. However, this debate has led us to call for a general transparency of processes, rather than for a specific set of criteria in this area.

### Discussion

The utility of qualitative research has been the subject of considerable debate. The tenor of this debate has frequently touched on the struggle to measure up to positivist constructs of what constitutes good research. For example, [Sandelowski and Barroso \(2002\)](#) note the apologetic stance some authors of qualitative studies take when describing small sample size as a limitation to the applicability of their findings. However, a preoccupation with generalisability, and thus with the quality criteria associated with this claim, represents a fundamental misunderstanding of the importance of qualitative research. The subjectivist epistemology upon which qualitative enquiry is based seeks to explore the 'how' and 'why' of human interactions, and is therefore communicating meanings and interpretations in the main. The strength of these approaches will be in understanding and explaining phenomena in similar settings. A number of factors, among them contextual, will shape how individuals respond to research findings. Qualitative research's strength has always been in illuminating context,

and our criteria reflect this central concern. Hence, there is significant emphasis on integrity, transparency and transferability in our appraisal checklist. In doing this, we make the explicit claim that some qualitative research can be inadequate.

An important decision regarding checklists is the degree of prescriptiveness with which they are applied. A number of voices have cautioned against a rigid model. [Barbour \(2001\)](#) argues for checklists to be viewed as 'reflective rather than constitutive of good research' (p. 1115), and specifically criticises the widespread uptake of theoretical sampling, grounded theory, multiple coding, triangulation and respondent validation as an unequivocal guarantee of robustness. These specific dimensions of qualitative enquiry need to be embedded within a broader understanding of qualitative research design and not 'stuck on as badge of merit' ([Barbour, 2001, p. 1115](#)). Similarly, [Sandelowski and Barroso \(2002\)](#) urge a flexible use of checklists, giving the benefit of the doubt to the researchers who, though they may have used inappropriate terminology in their papers, may still have produced worthwhile findings that can add to knowledge in the field. They distinguish between differing vehicles for publishing research, noting that journals usually require a template for recording that can truncate an original coherent report. It would be unfair to discredit a high-quality study on the basis of an imposed form required for publication. They make a plea for not confusing the actual research endeavour with the published form, and suggest reviewers 'read well' before judging whether the research was done well. Our position is that the checklist is indicative of quality but not a guarantee of it. We offer it for use imaginatively rather than prescriptively. For this reason, we are opposed to rating criteria and the production of a final quality score, as suggested by some authors ([Cesario et al., 2002](#)). As a baseline requirement, we would expect the criteria identified in column two of [Table 4](#) to be addressed adequately in qualitative research papers we include in a primary meta-synthesis. This may include the later discussion of insights gained from less well-conducted research at the end of a meta-synthesis review.

We acknowledge that we have arrived at the final checklist given in [Table 4](#) iteratively, rather than through a rigorous and systematic process. We also acknowledge that we are presenting yet another tool to add to the plethora of tools already in existence. However, we believe that, either through primary sources or through the three summary papers we examined, we have encountered most of the existing tools. We also believe

that ours is sufficiently compact to be of use to some researchers, specifically in the context of meta-synthesis, and specifically where a rapid review is needed. We are using the tool in our own review of qualitative synthesis of the nature of and phenomena within freestanding MLUs, and we have found it to be effective. In presenting the process of its development, our intention is to raise issues for debate as well as to offer one possible solution.

## Conclusion

Within health circles, interest in qualitative research is increasing. The trend is driven by the acknowledged complexity of many health-care interventions, the emphasis on client experience, and the focus on changing clinicians' practice. As the interest is translated into funding more studies, concern is being expressed about how to appraise these studies and, ultimately, what their findings mean for health-care practice. Although the literature on appraisal of qualitative research is extensive, there is clearly a lack of consensus on definitive criteria. The range of criteria exists on a continuum from endorsing positivist notions of reliability, validity and generalisability to a minimalist approach. Either end of the spectrum reflect different epistemological positions from positivism, with its dogma of objective truth, to postmodernist relativism for whom knowledge naming is like 'catching the wind.' This paper has attempted to ground criteria within the subjectivist epistemological position without embracing the multiplicity of little truths of more extreme postmodern positions. It values the personal, cultural and structural knowing that a variety of qualitative research methods aim to produce, while acknowledging that these may be reconstructed differently in the future. The criteria were synthesised iteratively from a variety of sources. We believe they form a working framework for qualitative research appraisal. We intend the tool to be applied reflexively and imaginatively. In that spirit, we hope the work will facilitate the identification of the strengths and limitations of papers from this research tradition.

## References

- Barbour, R., 2001. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *British Medical Journal* 322, 1117.
- Barroso, J., Gollop, C., Sandelowski, M., Meynell, J., Pearce, P., Collins, L., 2003. The challenge of searching for and retrieving qualitative studies. *Western Journal of Nursing Research* 25, 153–178.
- Bates, M., 1989. The design of browsing and berrypicking techniques for on-line search interface. *Online Review* 13, 407–431.
- Beck, C., 2002a. Mothering multiples: a meta-synthesis of qualitative research. *Maternal and Child Nursing* 27, 214–221.
- Beck, C., 2002b. Postpartum depression; a meta-synthesis. *Qualitative Health Research* 12, 453–472.
- Birch, M., Mauthner, M., Jessop, J., 2002. Introduction. In: Mauthner, M., Birch, M., Jessop, J., et al. (Eds.), *Ethics in qualitative research*. Sage Publications, London.
- Britten, N., Campbell, R., Pope, C., et al., 2002. Using meta ethnography to synthesize qualitative research: a worked example. *Journal of Health Services Research and Policy* 7, 209–215.
- Cesario, S., Morin, K., Santa-Donato, A., 2002. Evaluating the level of evidence in qualitative research. *Journal of Obstetrics, Gynaecology and Neonatal Nursing* 31, 531–537.
- Christians, C., 2000. Ethics and politics in qualitative research. In: Denzin, N., Lincoln, Y. (Eds.), *Handbook of qualitative research*. Sage, London.
- Clemmens, D., 2003. Adolescent motherhood: a meta-synthesis of qualitative studies. *American Journal of Maternal and Child Nursing* 28, 93–99.
- Clifford, J., Marcus, G., 1986. *Writing Culture: the poetics and politics of ethnography*. University of California Press, Berkeley.
- Critical Appraisal Skills Programme (CASP), 1999. Ten questions to help you make sense of qualitative research. CASP, Oxford.
- Crotty, M., 1998. *The foundations of social research: meaning and perspective in the research process*. London, Sage.
- Fox, N., 1993. *Postmodernism, sociology and health*. Open University Press, Buckingham.
- Glaser, B., Strauss, A., 1967. *The discovery of grounded theory*. Aldine, Chicago.
- Grol, R., 1997. Personal paper: beliefs and evidence in changing clinical practice. *British Medical Journal* 309, 1126–1128.
- Hammersley, M., Atkinson, P., 1995. *Ethnography: principles in practice*. Routledge, London.
- Harding, S., 1991. *Whose science? Whose knowledge? Thinking from women's lives*. Ithica, Cornell UP.
- Holloway, I., Wheeler, S., 1996. *Qualitative research for nurses*. Blackwell Science, Oxford.
- Harvey, L., 1990. *Critical Social Research*. Unwin Hyman, London.
- Hunt, S., Symonds, A., 1995. *The social meaning of midwifery*. MacMillan, Basingstoke.
- Kennedy, H., Rousseau, A., Low, L., 2003. An exploratory meta-synthesis of midwifery practice in the United States. *Midwifery* 19, 203–214.
- Kirkham, M., 1983. *Basic supportive care in labour; interaction with and around labouring women*. Unpublished PhD thesis. University of Manchester, Manchester.
- Kitzinger, S., 1962. *The experience of childbirth*. Victor Gollancz, London.
- Kuper, A., 1983. *Anthropology and anthropologists*. Routledge and Kegan Paul, London.
- Liamputtong, P., Naksook, C., 2003. Perceptions and experiences of motherhood, health and the husband's role among Thai women in Australia. *Midwifery* 19, 27–36.
- Mays, N., Pope, C., 2000. Qualitative research in health care: assessing quality in qualitative research. *British Medical Journal* 320, 50–52.

- Miles, M., Huberman, A., 1994. Qualitative data analysis: an expanded sourcebook. Sage, London.
- Morse, J., 1994. Emerging from the data: the cognitive processes of analysis in qualitative enquiry. In: Morse, J. (Ed.), *Critical issues in qualitative research methods*. Sage, London.
- Murphy, E., Dingwall, R., Greatbatch, D., et al., 1998. Qualitative research methods in health technology assessment: a review of the literature. DOH, York.
- Oakley, A., 1993. *Essays on women, medicine and health*. Edinburgh University Press, Edinburgh.
- Parsons, C., 1994. The impact of postmodernism on research methodology. *Nursing Inquiry* 2, 22–28.
- Paterson, B., Thorne, S., Canam, C., et al., 2001. *Meta-study of qualitative health research*. Sage, London.
- Perakyla, A., 1997. Reliability and validity in research based on transcripts. In: Silverman, D. (Ed.), *Qualitative research: theory, method and practice*. Sage, London.
- Popay, J., Rogers, A., 1998. Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research* 8, 341–352.
- Rolfe, G., 2000. Postmodernism: the challenge to empirical research. In: Rolfe, G. (Ed.), *Research, truth, authority: postmodern perspectives on nursing*. Macmillan, London, pp. 28–45.
- Spencer, L., Ritchie, J., Lewis, J., et al., 2003. Quality in qualitative evaluation: a framework for assessing research evidence. Government Chief Social Researcher's Office, Occasional Papers Series 2; <http://www.policyhub.gov.uk> (last accessed 5 July 2005).
- Sandelowski, M., Barroso, J., 2002. Reading qualitative studies. *International Journal of Qualitative Methods* 1, 1–47.
- Sandelowski, M., Docherty, S., Emden, C., 1997. Qualitative metasynthesis: issues and techniques. *Research in Nursing and Health* 20, 365–371.
- Strathern, M., 1988. *The gender of the gift*. University of California Press, Berkeley.
- Tobin, G., Begley, C., 2004. Methodological rigour within a qualitative framework. *Journal of Advanced Nursing* 48, 388–396.
- Walsh, D., Downe, S., 2005. Meta-synthesis method for qualitative research: a literature review. *Journal of Advanced Nursing* 50, 204–211.
- Yardley, L., 2000. Dilemmas in qualitative health research. *Psychology and Health* 15, 215–228.

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

