

## DEVELOPING RESEARCH

# The problem of appraising qualitative research

M Dixon-Woods, R L Shaw, S Agarwal, J A Smith

*Qual Saf Health Care* 2004;13:223–225. doi: 10.1136/qshc.2003.008714

Qualitative research can make a valuable contribution to the study of quality and safety in health care. Sound ways of appraising qualitative research are needed, but currently there are many different proposals with few signs of an emerging consensus. One problem has been the tendency to treat qualitative research as a unified field. We distinguish universal features of quality from those specific to methodology and offer a set of minimally prescriptive prompts to assist with the assessment of generic features of qualitative research. In using these, account will need to be taken of the particular method of data collection and methodological approach being used. There may be a need for appraisal criteria suited to the different methods of qualitative data collection and to different methodological approaches. These more specific criteria would help to distinguish fatal flaws from more minor errors in the design, conduct, and reporting of qualitative research. There will be difficulties in doing this because some aspects of qualitative research, particularly those relating to quality of insight and interpretation, will remain difficult to appraise and will rely largely on subjective judgement.

quality studies risk being flawed. For example, a study that tells us that junior doctors make errors primarily because of poor training might be wrong: the researchers might have relied on an inappropriate technique to investigate the issue or conducted the study badly, and it is possible that a better designed and conducted study could have come to a different conclusion—for example, emphasising issues of culture over those of training. Assessments of study quality are also needed for reviews. If studies of poor quality are included in a review, they may distort the synthesis in a range of possible ways and may cause difficulties in interpretation.<sup>10</sup>

### QUALITY CRITERIA FOR QUALITATIVE RESEARCH

Recent decades have seen the emergence of many proposals for quality criteria for qualitative research. These have been useful on the one hand in identifying criteria that might be used, but have also led to an unhelpful proliferation and diversity with often little evidence of common ground.<sup>11</sup> For example, it has been shown<sup>12</sup> that guidelines proposed by Seale and Silverman<sup>13</sup>—which emphasise the need for detailed transcription, the support of generalisations with counts of events, and the use of computer software—differ significantly from those proposed by Popay *et al*<sup>6</sup> which prioritise subjectivity, flexibility and adequate description and which argue that quasi-statistics and computer software are neither necessary nor sufficient for rigorous qualitative analysis. There are now over 100 sets of proposals on quality in qualitative research, some adopting non-reconcilable positions on a number of issues. For example, some recommend that qualitative studies should aim to be reproducible and that multiple coding is a good means of assessing the quality of qualitative research,<sup>14, 15</sup> while others<sup>16, 17</sup> deem such criteria meaningless when conducting research within a relativist paradigm involving multiple realities, subjectivity, and the negotiation of meaning. Attempts to produce consensus on criteria have proved difficult. The UK National Centre for Social Research has recently produced a framework for assessing qualitative research, drawing on 29 existing frameworks in the area as well as interviews with those active in the field and a workshop.<sup>18</sup> This has proved useful in describing the tensions and diversity in the field but it is lengthy—involving over 18 separate domains—and potentially unwieldy.

### QUALITATIVE RESEARCH IS NOT A UNIFIED FIELD

A key problem in much of the work on developing appraisal criteria for qualitative research has been the tendency to treat it as a

The valuable contribution that qualitative research can make to the study of quality and safety in health care is increasingly being recognised<sup>1</sup> and is well illustrated by several recent publications in *Quality and Safety in Health Care*.<sup>2, 3</sup> It is clearly important that policymakers and practitioners can have confidence in the quality of such research.<sup>4</sup> There is, however, disagreement not only about the characteristics that define good quality qualitative research, but also on whether criteria for quality in qualitative research should exist at all. Many argue that a set of criteria distinct from those applied to natural scientific quantitative approaches and specifically designed for qualitative research is required.<sup>4–6</sup> However, others have called for an end to “criteriology”,<sup>7</sup> arguing that this leads to privileging of method as a “sacred prescription” rooted in positivist philosophical traditions, and the stifling of the interpretive and creative aspects of qualitative research. Still others argue that criteria are best regarded as guides to good practice<sup>8</sup> rather than as rigid requirements in appraising papers.

Notwithstanding these debates, it is clear that some means of determining the quality of qualitative studies is needed.<sup>9</sup> Policy, practice, and clinical decisions made on the basis of low

See end of article for authors' affiliations

Correspondence to: M Dixon-Woods, Department of Health Sciences, University of Leicester, 22–28 Princess Road West, Leicester LE1 6TP, UK; md11@le.ac.uk

Accepted for publication 28 March 2004

unified field, both at the level of data collection (such as focus groups) and at the level of the methodological approach (such as grounded theory). Given the plurality of qualitative methodologies available, this is clearly a flawed approach destined to produce criteria that fit in certain cases but not in others. For example, Lincoln and Guba<sup>4</sup> and others recommend respondent validation where participants are asked to check researchers' interpretations. However, this may be an entirely unsuitable form of validity checking for some forms of qualitative research including, for example, discourse analysis, where its often anti-realist emphasis is on the multiple accounts that can be produced of any phenomenon rather than seeking a "single" verifiable account.

There is also a persistent failure to distinguish between the characteristics of a paper that are concerned with auditability and transparency of reporting, and those that are concerned with quality of process and analysis. This is not just a procedural problem or one of detail. It goes to the heart of what quality means in qualitative research. The dilemma is that some of the most important qualities of qualitative research can be the hardest to measure. For example, a study may be judged to have followed the appropriate procedures for a particular approach, to give information on selection of participants, and to provide clear details of the method followed. Yet the study may suffer from poor interpretation and offer little insight into the phenomenon at hand. On the other hand, a second study may be flawed in terms of the transparency of methodological procedures and yet offer a compelling, vivid and insightful narrative, grounded in the data. How the judgements of quality should be formed and informed in such cases is not at present clear. It does not follow from this that the project to address issues of quality in qualitative research should be abandoned—that is not helpful either intellectually or practically. However, it is crucial to take into account the particular features of qualitative research when considering the appropriate way of evaluating it.

### UNIVERSAL AND SPECIFIC: WE NEED BOTH

We argue that there are universal features of all forms of qualitative research, but these are few in number, may well be universal to all forms of research, and are best formulated as prompts to sensitise appraisers to the various dimensions of articles that require evaluation. The prompts in box 1 have been developed by a project team in the ESRC Research Methods Programme, following an empirical evaluation of using two existing frameworks and extensive discussions within a multidisciplinary team. In the development of these prompts we have been explicitly attentive to the need to avoid commitments to particular methodological approaches, and have distinguished between aspects of reporting and aspects of study design and execution. In addition, we wish to make clear that these are proposed as prompts to cue the attentiveness to a specific set of issues. They are not criteria, do not prescribe how the reader makes the assessment, and do recognise that the assessment will inevitably involve subjective judgement on the part of the reader. Because these prompts are generic, it would also be possible for them to be complemented by prompts that are specific to different methods of data collection and qualitative methodologies. So far there have been few attempts to develop such methodology-specific approaches.

The need for such approaches is demonstrated by even the most cursory overview of currently published qualitative work in a wide range of journals. It is often difficult even to be sure about which study design—let alone which theoretical perspective—a published study has used. Researchers may describe something as qualitative research when, in the judgement of others, it is not. Similarly, studies may be

### Box 1 Prompts for appraising qualitative research

- Are the research questions clear?
- Are the research questions suited to qualitative inquiry?
- Are the following clearly described?
  - sampling
  - data collection
  - analysis
- Are the following appropriate to the research question?
  - sampling
  - data collection
  - analysis
- Are the claims made supported by sufficient evidence?
- Are the data, interpretations, and conclusions clearly integrated?
- Does the paper make a useful contribution?

described as using a specific approach such as grounded theory when, in the judgement of others, it has not been used.<sup>19</sup>

Indeed, grounded theory can be seen as an example of a particularly problematic case. Here the original text by Glaser and Strauss<sup>20</sup> is one of the most frequently cited social science texts. However, the initial vision is more frequently honoured in the breach than the observance.<sup>21</sup> Inconsistencies, misappropriations, and mislabelling of studies purporting to use grounded theory are common.<sup>22</sup> Some of the reasons for this lie in the practical difficulties of implementing grounded theory in its original form.<sup>23</sup> The problem is further exacerbated by the fissure that arose between Glaser and Strauss themselves, who went on to develop two quite distinctive approaches to grounded theory. Strauss's later work is focused on the development of technique and procedure and could be seen as rather prescriptive in character, while Glaser's work can appear to be lacking in a grounding in practicalities and has largely avoided the development of procedural techniques. It is, however, clear that some elements of grounded theory (particularly the constant comparative method of analysis) are especially useful, but the tendency to select some of these techniques to create ad hoc and "à la carte" approaches to qualitative analysis and still retain the label "grounded theory" is very unhelpful. The result is that, despite the tendency for researchers to assume that describing a study as using "grounded theory" is sufficient, it is far from self-evident what such a study might involve.

The situation in qualitative research contrasts with the development of methods for appraising quantitative research where it is recognised (in the work of the Critical Appraisal Skills Programme (CASP)<sup>24</sup> and the NHS Centre for Reviews and Dissemination (CRD) guidance,<sup>11</sup> for example) that different study types (such as randomised controlled trials, case control studies, studies to evaluate screening programmes) may demand different criteria. This allows the precise formulation of flaws that would be fatal or very damaging to the rigour of a particular study type. For example, failure to randomise properly in a randomised controlled trial, or to select appropriate controls in a case control study, could be deemed very serious problems. In

### Key points

- Some means of appraising qualitative research is needed if it is to contribute appropriately to systematic reviews.
- Proposals for criteria that might define high quality qualitative research have proliferated, but sometimes do not overlap or are difficult to operationalise.
- A minimal set of prompts is proposed to help cue attention to the range of dimensions of qualitative research that require appraisal.
- There is a need for additional criteria that recognise the diversity of study designs and theoretical perspectives in qualitative research, and to distinguish between minor errors and fatal flaws.
- The measurement of all aspects of quality of qualitative research will remain difficult.

qualitative research there is a need to recognise that focus groups, interviews, participant observation, and so on also constitute different study types and therefore may also need different criteria to allow their appraisal, and similarly have different types of fatal flaws. It is also important to recognise that the classification of minor errors and fatal flaws may not, however, be easily distinguished into simple binary judgements: it may be necessary to take a holistic view of a study that recognises the importance of context and what was feasible in that context.

It is also vitally important to recognise that, to a much greater extent than in quantitative research, the execution of a qualitative research study type is crucially related to the *theoretical perspective* in which the researchers have chosen to locate the study. An interview based study conducted within a grounded theory framework may therefore have very different characteristics from an interview based study conducted within a discourse analysis framework, and this clearly needs to be reflected in the framework for appraising a study. The key task then lies in defining what the expectations should be for a particular study design within a particular theoretical field. Unless we can move to these kinds of definitions and expectations, the diversity—indeed, near anarchy—in qualitative methodology means that it is very difficult to identify, or at least gain agreement, on what might constitute a fatal flaw in, for example, an interview based study using an opportunistic sample and the constant comparative method for analysis.

### CONCLUSIONS

Qualitative research has an important and growing role in the study of quality and safety. The rationale for establishing the quality of qualitative research is clear. Problems occur when attempting to determine exactly how this task can proceed. It is evident that debate and discussion, together with systematic evaluation of the implications of adopting various methods for appraisal, are necessary to resolve many of the critical issues raised here. We have proposed a minimal set of prompts which have been designed to stimulate appraisal of different dimensions of qualitative research but are explicitly methodology neutral. We argue that there is a need for future development in this area to focus on the distinctive study designs and theoretical perspectives that qualitative research

can adopt, to distinguish fatal flaws from more minor errors, and to recognise that many of the more important and interesting aspects of qualitative research may remain very difficult to measure except through the subjective judgement of experienced qualitative researchers.

### ACKNOWLEDGEMENTS

The authors thank the Health Development Agency and the ESRC Research Methods Programme (award number H333250043) for funding the work on which this paper is based, Sheila Bonas, Tina Miller and Bridget Young for their contribution to the development of the prompts, and David Jones and Alex Sutton for helpful comments on earlier drafts.

### Authors' affiliations

**M Dixon-Woods, R L Shaw, S Agarwal**, Department of Health Sciences, University of Leicester, Leicester LE1 6TP, UK

**J A Smith**, Department of Psychology, Birkbeck College, University of London, London WC1E 7HX, UK

Conflicts of interest: none declared.

### REFERENCES

- 1 **Pope C**, Van Royen P, Baker R. Qualitative methods in research on healthcare quality. *Qual Saf Health Care* 2002;**11**:148–52.
- 2 **Taxis K**, Barber N. Causes of intravenous medication errors: an ethnographic study. *Qual Saf Health Care* 2003;**12**:343–7.
- 3 **Bradley EH**, Holmboe ES, Maittey JA, et al. Data feedback efforts in quality improvement: lessons learned from US hospitals. *Qual Saf Health Care* 2004;**13**:26–31.
- 4 **Lincoln YS**, Guba EG. *Naturalistic inquiry*. Newbury Park, CA: Sage, 1985.
- 5 **Seale C**. Quality issues in qualitative inquiry. *Qualitative Social Work* 2002;**1**:97–110.
- 6 **Papay J**, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. *Qual Health Res* 1998;**8**:341–51.
- 7 **Schwandt TA**. Farewell to criteriology. *Qualitative Inquiry* 1996;**2**:58–72.
- 8 **Henwood K**, Pidgeon N. Qualitative research and psychological theorising. *Br J Psychol* 1992;**83**:97–111.
- 9 **Dixon-Woods M**, Fitzpatrick R. Qualitative research in systematic reviews has established a place for itself. *BMJ* 2001;**323**:765–6.
- 10 **Campbell R**, Pound P, Pope C, et al. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Soc Sci Med* 2003;**56**:671–84.
- 11 **NHS Centre for Reviews and Dissemination**. *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews*, 2nd ed. Report Number 4. York: NHS CRD, 2001.
- 12 **Chapple A**, Rogers A. Explicit guidelines for qualitative research: a step in the right direction, a defence of the 'soft' option, or a form of sociological imperialism? *Fam Pract* 1998;**15**:556–61.
- 13 **Seale C**, Silverman D. Ensuring rigour in qualitative research. *Eur J Public Health* 1997;**7**:379–84.
- 14 **Engel JD**, Kuzel AJ. On the idea of what constitutes good qualitative inquiry. *Qual Health Res* 1992;**2**:504–10.
- 15 **Strauss A**, Corbin J. *Basics of qualitative research: techniques and procedures for developing grounded theory*, 2nd ed. Thousand Oaks, CA: Sage, 1998.
- 16 **Yardley L**. Dilemmas in qualitative health research. *Psychol Health* 2000;**15**:215–28.
- 17 **Morse JM**. *Critical issues in qualitative research methods*. Thousand Oaks, CA: Sage, 1994.
- 18 **Spencer I**, Ritchie J, Lewis J, et al. *Quality in qualitative evaluation: a framework for assessing research evidence*. Government Chief Social Researcher's Office, Occasional Series No 2, 2003, [http://policyhub.cmps.gov.uk/servlet/DocViewer/doc=1421/qual\\_eval\\_exe\\_summ.htm](http://policyhub.cmps.gov.uk/servlet/DocViewer/doc=1421/qual_eval_exe_summ.htm) (accessed 21 September 2003).
- 19 **Paterson BL**, Thorne SE, Canam C, et al. *Meta-study of qualitative health research: a practical guide to meta-analysis and meta-synthesis*. Thousand Oaks, CA: Sage, 2001.
- 20 **Glaser B**, Strauss A. *The discovery of grounded theory: strategies for qualitative research*. Chicago: Aldine, 1967.
- 21 **Bryman A**, Burgess RG, eds. *Analysing qualitative data*. London: Routledge, 1994.
- 22 **Stern PN**. Eroding grounded theory. In: Morse JM, ed. *Critical issues in qualitative research methods*. Thousand Oaks, CA: Sage, 1994.
- 23 **Barbour R**. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *BMJ* 2001;**322**:1115–7.
- 24 **CASP**. <http://www.phru.org.uk/~casp/casp.htm> (accessed 21 September 2003).