Journal of Evaluation in Clinical Practice, **9**, 2, 187–194

Reframing the evaluation of qualitative health research: reflections on a review of appraisal guidelines in the health sciences

Joan M. Eakin PhD[1] and Eric Mykhalovskiy PhD[2]
[1]Associate Professor, Public Health Sciences, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
[2]Assistant Professor, Department of Community Health and Epidemiology, Dalhousie University, Halifax, NS, Canada

Correspondence
Dr Joan M. Eakin
Department of Public Health Sciences
Faculty of Medicine
University of Toronto
McMurrich Building
12 Queen's Park Crescent West
Toronto, ON M5S 1A8
Canada
E-mail: joan.eakin@utoronto.ca

Keywords: critique, evaluation criteria, health sciences, qualitative methods, reformulation

Accepted for publication:
21 October 2002

Abstract

In this article, we explore the form of evaluation put forward by guidelines used in the health sciences for appraising qualitative research and we begin to articulate an alternative posture. Most guidelines are derivative of the modes of assessment developed by clinical epidemiologists as part of the promotion of evidence-based medicine (EBM). They are predominantly proceduralist in orientation, equating quality with the proper execution of research techniques. We argue that this form of judgment assumes a fixed relationship between research practice and knowledge generated, and tends to over-simplify and standardize the complex and non-formulaic nature of qualitative inquiry. A concern with methods as objects of judgment in and of themselves restricts the reader's field of vision to the research process and diverts attention away from the analytic content of the research. We propose an alternative 'substantive' perspective that focuses on the analysis put forward, and regards methods as resources for engaging with and understanding the substantive findings and topic of inquiry. An important challenge is to find a way to embody such a form of judgment in practical assessment tools.

Introduction

In recent years, qualitative research has established an important foothold in the health sciences. Qualitative research exploring clinical decision making, practitioner–patient interaction, patient experiences of illness and other topics is now published routinely in medical, nursing and other health journals. A host of journal articles introducing the central features of qualitative research to health practitioners and researchers has also been published (Berkwits & Aronowitz 1995; Black 1994; Britten *et al.* 1995; Goering & Streiner 1996; Miller & Crabtree 1994; Needleman & Needleman 1996; Pope & Mays 1995; Pope & Ziebland 2000; Sofaer 1999; Spencer 1993; Willms *et al*. 1997).

The growing popularity of qualitative research in the health sciences has occasioned a crisis of evaluation. Good qualitative research is methodologically and epistemologically distinct from the clinical and epidemiological sciences with which clinicians and other members of the health sciences community are most familiar. Qualitative research that is largely narrative, that relies fundamentally on language-based data and whose relevance is not secured through a numerical calculus of 'confidence intervals', '*P* values' and the like presents problems for the uninitiated reader. In response, a literature has developed around the question of how best to understand, evaluate and apply qualitative research in the health sciences. This literature typically draws on and modifies critical appraisal and related approaches to evaluation that were initially formulated by clinical epidemiologists and health services researchers as part of the promotion of evidence-based decision making in health care. Thus, the literature offers

readers a variety of guidelines that structure evaluation through a standard set of criteria of quality, often in the form of a checklist of questions to be posed of various aspects and stages of the research process.

In health science circles, assessment guidelines of this sort have appeal because they offer a practical and efficient means of steering journal and funding agency review processes and of instructing readers in how to appraise the qualitative literature they read. As such, they are powerful textual devices that have the potential to substantially influence the conduct and reception of qualitative health research. Despite their proliferation and significance for shaping the way in which qualitative inquiry is evaluated and used in the health sciences, these assessment devices have received scant scrutiny.

This paper extends recent discussions of the relationship of qualitative research to evidence-based medicine (EBM) (Dixon-Woods, Fitzpatrick & Roberts 2001; Upshur, VanDenKerkhof & Goel 2001) through a focus on evaluation. We propose that EBM-derived criteria checklists direct readers into a form of judgment that is not well suited to the character of qualitative research, and that limits their capacity to relate to – and ultimately assess, appreciate and use – the distinctive contributions of this kind of research. We base our critique on a review of a sample of qualitative assessment guidelines published in health science journals, and we put forward suggestions for an alternative basis for judging qualitative research.

We focused our review on published guidelines in health science journals that appear to be aimed at relatively novice readers, and that formulate explicit considerations for the judgment of qualitative research. We examined many guidelines, but concentrated detailed analysis on a sample of articles that represented a diversity of health disciplines, and that were published in well-known journals or cited often in the qualitative health research assessment literature. Our sample included Burns (1988), Cobb & Hagemaster (1987), Dowell, Huby & Smith (1995), Elder & Miller (1995), Giacomini, Cook & Group (2000), Greenhalgh & Taylor (1997), Inui (1991), Krefting (1990), Kuzel *et al.* (1994), Leininger (1994), Mays & Pope (1995, 2000a), Patton (1999), Rowan & Huston (1997) and Ward (1993). Although they offered varying amounts of supporting explanation and elaboration, all the articles attempted to outline specific criteria for assessing qualitative research reports, usually in question format, as illustrated in Box 1.

Our examination of the guidelines was conducted as a text analysis of the criteria used and the rationale offered for their relevance and use. We focused on criteria related to the features of the research process considered key in the guidelines themselves: research question; design; role of the researcher; sampling; data collection, and analysis. For each feature, in each guideline, we generated analytic summaries of the kinds of questions readers are urged to ask, the explicit assessment criteria proposed and their explicit and implicit assumptions and logic.

Our observations were informed by editorial commentary on the assessment of qualitative health research (Gilchrist & Engel 1995; Popay 1995; Poses & Isen 1998), by discussions of the distinct epistemological and methodological features of qualitative inquiry in the health field (Barbour 2001; Devers 1999; Emden & Sandelowski 1998, 1999) and of its desirable features (Dingwall *et al.* 1998; Harding & Gantley 1998; Popay, Rogers & Williams 1998) and by general discussions of research quality and validity (Guba & Lincoln 1989; Hammersley 1998; Johnson, Long & White 2001; Seale 1999; Smith & Deemes 2000; Sparkes 2001; Whittemore, Chase & Mandle 2001).

Box 1 Examples of guideline assessment questions

- 'Are the data used appropriate to the question?' (Inui 1991, p. 485)
- 'Did the sample include the full range of possible cases or settings?' (Mays & Pope 2000, p. 52)
- 'What methods did the researcher use for collecting data – and are these described in enough detail?' (Greenhalgh & Taylor 1997, p. 741)
- 'Are the unique issues of sampling in a qualitative study adequately addressed?' (Cobb & Hagemaster 1987, p. 141)
- 'Did the investigators become too close to their participants?' (Elder & Miller 1995, p. 293)
- 'Does the analysis involve interpretation, not merely counting the frequency of events or categories?' (Dowell *et al.* 1995, p. 46)

Procedural judgment

Our central observation is that the guidelines advance a form of judgment that is primarily proceduralist in nature. Firstly, the actual form of the guidelines is procedural: standardized checklists of criteria for particular 'stages' of the research process invite little more than 'yes' or 'no' answers. Secondly, the content of the guidelines is procedural: qualitative studies are to be appraised in terms of how well they were conducted according to particular standards of research practice. That is, quality is seen to arise from how the research was carried out – from the proper execution of scientific method.

The procedure-centred nature of assessment is evident across all areas of research process. In relation to the research question, the notion of a 'qualitative question' (Kuzel *et al*. 1994) is implicit in many texts. Readers are prompted to ask if the research question is of the 'right' sort for qualitative inquiry, with 'right' questions being those that coincide with the core features and applications of qualitative research, such as 'deeper understanding' (Greenhalgh & Taylor 1997, p. 742), exploration and description (Giacomini *et al*. 2000), a sensitivity to context (Popay *et al*. 1998) and, most commonly, the study of subjective meaning and experience (Cobb & Hagemaster 1987; Greenhalgh & Taylor 1997; Inui 1991; Popay *et al*. 1998). The criterion is thus a 'match' between question and method.

Procedural criteria are also used to appraise the role of the researcher, a second area considered key to quality assessment. Many checklists acknowledge qualitative research as inherently and unavoidably 'subjective', and consider the researcher's active presence in the research as a source of 'bias' that needs to be addressed in order to keep the research 'relatively untainted with personal or cultural perspective' (Greenhalgh & Taylor 1997, p. 742). Guidelines often call for a declaration of the researcher's perspective or indicate that it should be made visible to readers (e.g. Inui 1991; Rowan & Huston 1997). Although some writers see the investigator as an important 'instrument' of observation and research (e.g. Inui & Frankel 1991, Mays & Pope 2000a), or as a source of knowledge (Krefting 1990), there is a pervasive and enduring ambivalence in many guidelines towards the role of the researcher. For the most part, self-reflexivity is valued as a means to avert problems of bias, the 'over-enmeshment' of researchers with participants (Elder & Miller 1995, p. 283) and other threats to objectivity, thereby increasing the credibility of the research (Krefting 1990).

Criteria for assessing sampling also centre on a concern with method. For example, readers are enjoined to ask whether the sampling process has been described clearly, if 'sound reasoning' (Giacomini *et al*. 2000, p. 358) is provided for its justification, if the sampling strategy matches the purpose of the research and methods used (Elder & Miller 1995; Greenhalgh & Taylor 1997; Huston & Rowan 1997) and if sampling procedures have maximized the 'richness' and 'diversity' of data (Dowell *et al*. 1995; Elder & Miller 1995; Krefting 1990). 'Convenience' sampling is widely considered to be less acceptable (Elder & Miller 1995; Giacomini *et al*. 2000; Mays & Pope 1995).

Procedural criteria also prevail in the consideration of data collection. Although most of the guidelines are sensitive to the unique iterative relationship between qualitative data collection and analysis, most criteria focus on the transparency of the study's data collection process. Most guidelines urge readers to look for specification of the type of data collection strategy, the rationale for the choice and details on the structure, content and circumstances of data collection. Readers are advised that 'researchers need to choose the [data collection] method best suited for their study' (Elder & Miller 1995, p. 281) and are urged to ensure that strategies are 'congruent with the purpose of the study' (Cobb & Hagemaster 1987, p. 144). Criteria tend to privilege triangulation – the use of more than one method of data collection.

Assessment of the analysis component of research also relies heavily on indicators of procedural appropriateness. Guidelines urge readers to seek specification of procedures used, such as cross-coding, constant comparison, discrepant case analysis, triangulation, saturation and audit trails (e.g. Krefting 1990; Mays & Pope 1995; Ward 1993). Validity is to be ascertained on the basis of the adequacy of the evidence, such as 'thickness' of the description, the proper use of supporting data and the logical consistency of the argumentation (e.g. Greenhalgh & Taylor 1997; Huston & Rowan 1997).

Thus, in relation to various parameters of the research process, the guidelines in our sample adopt a strongly procedure-oriented approach: quality is to be gauged on the basis of the researcher having made the right choice of method and having executed it in the right way. In our view, however, there are a number of reasons why this form of judgment is inadequate for apprehending qualitative research.

Firstly, in the case of critical appraisal, the standardized checklist approach to assessment works in large part because most readers – practitioners as well as researchers – are familiar with the basic underlying logic and form of scientific inquiry under consideration. In the case of qualitative research, while novice health science readers may have some awareness of its techniques (e.g. focus groups, identifying 'themes'), they often do not share an understanding of its fundamental principles and type of reasoning (e.g. nature of meaning, forms of generalization), which is needed to apply the criteria or answer the questions posed in the checklists. For example, to assess if the context has been 'adequately described' or if 'sound reasoning' has been provided for choice of sampling strategy, the reader needs to understand the multiple roles of context in the qualitative research process, and how sampling is used strategically in qualitative inquiry to facilitate opportunities for analytic comparison.

A second concern with a proceduralist approach stems from problems with the procedures themselves, particularly those modeled too closely on principles of positivist science. An example is the treatment of the research question as an autonomous place of genesis from which the subsequent research process unfolds (reflecting the maxim that the method depends on the question being asked). In many forms of qualitative inquiry, however, the research question functions more as a compass than as an anchor, and is sometimes not really known until the end of the research.

Another example of inappropriate transfer of scientific conceptions of quality lies in the value accorded to triangulation of data collection, based on the belief that multiple data sources give a more valid picture of reality. From a qualitative perspective, triangulation can be seen as producing not different dimensions of the same thing but different things (Holland & Gillies 2002; Kvale 1996). Further imports from positivist methodology include the treatment of 'themes' as relatively objective phenomena inherent in the data that 'emerge' or self-identify, and the notion that interpretations can be 'verified' through cross-coding by colleagues or subjects. Guidelines that urge readers to assess quality on the basis of such procedures can lead to in inappropriate evaluation of the qualitative research process.

A third limitation of a procedural approach is that it is restrictive, and unable to manage the non-formulaic complexity of the qualitative research process. For example, too standardized a set of prescriptions for assessing analysis may prevent readers from recognizing and appreciating fertile forms and sites of analytic activity, such as overlooking or discrediting evidence that is gleaned from the ensemble of the data rather than from discrete, identifiable segments, or evidence that resides 'between the lines' or in silence. A technical approach to assessment can also fail to capture the complex balance between over-theorization (the imposition of a theoretical straightjacket on the interpretation) and under-theorization (the failure to conceptualize the data beyond everyday 'common sense'). Such judgment rests more on a gestalt assessment of the researcher's theoretical stance and on the interface between the reader's and the writer's interpretive platforms – attributes not captured easily in a standardized checklist of analytic procedures.

Proceduralist assessment also provides a restrictive basis for judging what some perceive to be the central analytic enterprise of qualitative inquiry: writing practices (Richardson 2000). Researchers articulate their standpoint and their substantive findings through the words they use and the overall rhetorical form of their writing. Hence, the analysis cannot be fully appraised apart from its written character. This feature of qualitative analysis, if it is addressed at all in guidelines, is presented in terms of the writing style or technique, which glosses over more complex underlying issues of representation (Atkinson 1990; Fine *et al.* 2000).

The fourth and most important problem with procedural judgment is that it absorbs so much evaluative energy that almost no attention is focused on the substantive offering of the research – on what the authors actually say about the phenomena they have

investigated and how they relate their research practices to their findings. Let us illustrate what a 'substantive' approach to assessment might mean in relation to the research stages we have discussed.

Substantive judgment

A substantive approach is one in which the reader's evaluative gaze is broadened from a focus on proper method to a focus on analytic content. This does not mean that research process is unimportant, only that it ceases to be an object of judgment in and of itself, and becomes relevant as an aid to the interpretation of and engagement with the substantive findings and topic of inquiry.

With respect to the research question, rather than fore-fronting the issue of whether or not a question is a 'correct' choice for a qualitative approach, a 'substantive' form of judgment might urge readers to use the research question as a positioning device for apprehending the nature of the investigation and, particularly, for understanding its findings. Here, the research question would orient readers to the kinds of knowledge that were sought in the research process, would inform their appreciation of the interpretative stance taken to the data and would constitute a point of reference for comprehending the substantive analysis put forward. A sense of how and why the research question changed over the course of the research can engage the reader in the research process and in the ultimate analysis in a way not invited by a formulaic 'right–wrong' assessment of the research question.

With regard to the role of the researcher, a substantive approach would be based on an understanding of researcher subjectivity not as a problem of bias to be eliminated or reduced (or at least confessed) but as something to be used actively and creatively throughout the research process (Hertz 1997; Krieger 1991; Lincoln & Guba 2000). Judgment might begin from a place that fully recognizes subjective knowledge as a *resource*, that treats it as a critical point of reference that allows readers to apprehend both the research process and the substantive analysis being proposed. This evaluative stance 'frees up' the role of the researcher from a narrow methodologism and invites readers to explore what researchers say about their role and involvement in their research in ways that depart from a mooring in a more or less explicit call for neutrality. Readers can reflect on how the class, gender and other features of the researcher's social location bear upon the kind of analysis and interpretation of findings that are made. Of course, there are unlimited aspects to the researcher's relationship to the data, which places the onus on the researcher to choose and justify which features of the relationship are relevant and to describe their role in data collection and analysis.

Similarly, with regard to sampling, a substantive approach would have readers inquire into how the sampling process and the sample itself bear on the data collected and on the interpretations made of them. Here, the sampling process becomes a window on other elements of the research process rather than just an indicator of generalizability. For example, information on the characteristics of participants or the organizational setting is useful not just for what it says about the validity of findings for other types of people or situations, but also for the light it casts on the circumstances that shaped interaction during data gathering and the data obtained. It contributes to the reader's ability to 'make sense' of how the research was done and the claims being made about it, as well as how and why the analytic focus evolved the way it did.

In a parallel way, information on how data were collected is important, not just to demonstrate that the research was systematic (a hallmark of science), but also to enhance the reader's grasp of the meaning being attributed to the data. For example, knowing that one-on-one oral interviews were used to gather data from persons of low educational and social status might make the reader reflect on what might not have been said in such circumstances due to social unease with the form of interaction.

Regarding the analysis stage of research, judgment based on a search for documentation of the rules and procedures of analytic technique is not very productive because there is no direct relationship between specific techniques and specific analytic outcomes. Although procedures can function as important triggers to analytic insight, they do not themselves constitute that insight. A substantive approach would instead try to enhance the readers' capacity to 'feel' the texture of the account being put forward, to

understand the conceptual development and foundations of the analysis and thereby, ultimately, to better apprehend the leaps of imagination and creative thinking that constitute (arguably) the most valuable feature of all research, qualitative or otherwise.

Conclusion

In this article, we have critiqued the form of judgment put forward by guidelines used in the health sciences for appraising qualitative research and have begun to articulate an alternative evaluative posture. We have argued that by gauging quality primarily in terms of the execution of proper method, the guidelines tend to over-simplify and inappropriately standardize the complex and non-formulaic nature of qualitative inquiry, promoting the notion of a fixed relationship between research practice and knowledge generated. More importantly, the guidelines' concern with procedural correctness restricts the reader's field of vision on the research process and diverts attention away from the analytic content of the research.

As an alternative, we propose augmenting procedural considerations with a 'substantive' orientation that centres on the relationship between research practices and substantive findings and interpretation. Evaluative schema would not simply ask about which research procedures were used, whether they were the correct ones to use and whether they testify to the validity of the findings. Rather, judgment would emerge from a deeper engagement in and understanding of the interpretations and propositions being put forward and assessment of how they are (or are not) produced and rendered convincing by the research practices used. From this perspective, procedural considerations function less as autonomous indicators of 'good' or 'bad' quality and more as resources for 'making sense' of the analysis.

However, how such a quintessentially qualitative form of judgment can be incorporated into an assessment tool remains a formidable challenge. How much standardization is possible without rendering the evaluative exercise superficial or compromising core methodological principles of interpretive inquiry? What are alternative formats to checklists and other formulaic structures? How can the complexity of qualitative research be written intelligibly and succinctly for relatively novice readers without 'dumbing down' the methodological and theoretical issues?

It is important to find answers to such questions because, in the absence of alternative formulas, journal editors, reviewers and readers will continue to use existing tools, and will continue to be underserved by them. The need to revise current guidelines is also underscored by the likelihood that qualitative researchers, who cannot ignore the forms of judgment brought upon their work when they seek funding or publication, will find themselves 'writing to' the procedural considerations embodied in the criteria or, even worse, actually doing the research in this way.

Nevertheless, in the health sciences, it will not be easy to shift away from dominant models of scientific appraisal. Asserting a different form of judgment (as opposed to just fine-tuning existing models) is discouraged by the marginality of qualitative inquiry in the health research field as a whole, and by the associated quest of its practitioners for scientific legitimacy, professional acceptance and publication success. This situation is compounded by 'tensions between the pragmatic world of health care management and the quality and integrity of qualitative research' (Peck & Seeker 1999, p. 552), the 'problem-centred' nature of health services research (Harding & Gantley 1998) and the search of health sciences journals for easy-to-use, one-size-fits-all templates for reviewing qualitative research.

References

Atkinson P. (1990) *The Ethnographic Imagination: Textual Constructions of Reality.* Routledge, London.

Barbour R.S. (2001) Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? *British Medical Journal* **332** (7294), 1115–1117.

Berkwits M. & Aronowitz R. (1995) Different questions beg different methods. *Journal of General Internal Medicine* **19**, 409–410.

Black R. (1994) Why we need qualitative research. *Journal of Epidemiology and Community Health* **48**, 425–426.

Britten N., Jones R., Murphy E. & Stacy R. (1995) Qualitative research methods in general practice and primary care. *Family Practice* **12** (1), 104–114.

Burns N. (1988) Standards for qualitative research. *Nursing Science Quarterly Nursing Science Quarterly*, 44–52.

Cobb A.K. & Hagemaster J.N. (1987) Ten criteria for evaluating qualitative research proposals. *Journal of Nursing Education* **26** (4), 138–143.

Devers K.J. (1999) How will we know 'good' qualitative research when we see it? Beginning the dialogue in health services research. *Health Services Research* **34** (5), 1153–1188.

Dingwall R., Murphy E., Watson P., Greatbatch D. & Parker S. (1998) Catching goldfish: quality in qualitative research. *Journal of Health Services Research Policy* **3** (3), 167–172.

Dixon-Woods M., Fitzpatrick R. & Roberts K. (2001) Including qualitative research in systematic reviews: opportunities and problems. *Journal of Evaluation in Clinical Practice* **7** (2), 125–134.

Dowell J., Huby G. & Smith C. (1995) *Scottish Consensus: Statement on Qualitative Research in Primary Health Care.* Tayside Centre for General Practice, University of Dundee, Dundee.

Elder N. & Miller W. (1995) Reading and evaluating qualitative research studies. *The Journal of Family Practice* **41** (3), 279–285.

Emden C. & Sandelowski M. (1998) The good, the bad and the relative, part one: conceptions of goodness in qualitative research. *International Journal of Nursing Practice* **4** (4), 206–212.

Emden C. & Sandelowski M. (1999) The good, the bad and the relative, part two: goodness and the criterion problem in qualitative research. *International Journal of Nursing Practice* **5** (1), 2–7.

Fine M., Weis L., Weseen S. & Wong L. (2000) For whom? Qualitative research, representations and social responsibilities. In *Handbook of Qualitative Research*, 2nd edn (eds N. Denzin & Y. Lincoln), pp. 107–131. Sage Publications, Thousand Oaks, New Delhi, London.

Giacomini M.K. & Cook D.J. (2000) Users' guides to the medical literature, XXIII. Qualitative research in health care, A. Are the results of the study valid? *Journal of the American Medical Association* **284**, 357–362.

Gilchrist V.J. & Engel J.D. (1995) Qualitative research and clinical care. *The Journal of Family Practice* **41** (3), 229–230.

Goering P. & Streiner D. (1996) Reconcilable differences: the marriage of qualitative and quantitative methods. *Canadian Journal of Psychiatry* **41**, 491–497.

Greenhalgh T. & Taylor R. (1997) Papers that go beyond numbers (qualitative research). *British Medical Journal* **315**, 740–743.

Guba E. & Lincoln Y. (1989) *Fourth Generation Evaluation*. Sage Publications, Newbury Park, CA.

Hammersley M. (1998) *Reading Ethnographic Research: a Critical Guide.* Longman, London.

Harding G. & Gantley M. (1998) Qualitative methods: beyond the cookbook. *Family Practice* **15**, 76–79.

Hertz R. (1997) *Reflexivity and Voice.* Sage Publications, Thousand Oaks, CA.

Huston P. & Rowan M. (1997) Qualitative research articles: information for authors and peer reviewers. *Canadian Medical Association Journal* **157** (10), 1442–1446.

Inui T. & Frankel R. (1991) Evaluating the quality of qualitative research. *Journal of General Internal Medicine* **6**, 485–486.

Johnson M., Long T. & White A. (2001) Arguments for 'British pluralism' in qualitative health research. *Journal of Advanced Nursing* **33** (2), 243–249.

Krefting L. (1990) Rigour in qualitative research: the assessment of trustworthiness. *American Journal of Occupational Therapy* **45** (3), 214–222.

Krieger S. (1991) *Social Science and the Self: Personal Essays on an Art Form.* Rutgers University Press, New Brunswick, NJ.

Kuzel A., Engel J., Addison R. & Bogdewic S. (1994) Desirable features of qualitative research. *Family Practice Research Journal* **14** (4), 369–378.

Kvale S. (1996) *Interviews: an Introduction to Qualitative Research Interviewing.* Sage Publications, Thousand Oaks, CA.

Leininger M. (1994) Evaluation criteria and critique of qualitative research studies. In *Critical Issues in Qualitative Research Methods* (ed. J. Morse), pp. 95–115. Sage Publications, Thousand Oaks, London, New Delhi.

Lincoln Y.S. & Guba E.G. (2000) Paradigmatic controversies, contradictions, and emerging confluences. In *Handbook of Qualitative Research*, **Vol. 2** (eds N.K. Denzin & Y.S. Lincoln), pp. 163–188. Sage Publications, Thousand Oaks, CA.

Mays N. & Pope C. (1995) Rigour and qualitative research. *British Medical Journal* **311**, 109–112.

Mays N. & Pope C. (2000a) Assessing quality in qualitative research. *British Medical Journal* **320**, 50–52.

Miller W. & Crabtree B. (1994) Qualitative analysis: how to begin making sense. *Family Practice Research Journal* **14** (3), 289–297.

Needleman C. & Needleman M. (1996) Qualitative methods for intervention research. *American Journal of Industrial Medicine* **29** (4), 329–337.

Patton M.Q. (1999) Enhancing the quality and credibility of qualitative analysis. *Health Services Research* **34** (5), 1189–1208.

Peck E. & Seeker E. (1999) Quality criteria for qualitative research: does context make a difference? *Qualitative Health Research* **9** (4), 552–558.

Popay J. (1995) Qualitative research and the gingerbread man. *Health Education Journal* **54**, 389–392.

Popay J., Rogers A. & Williams G. (1998) Rationale and standards for the systematic review of qualitative literature in health services research. *Qualitative Health Research* **8** (3), 341–351.

Pope C. & Mays N. (1995) Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. *British Medical Journal* **311**, 42–45.

Pope C. & Ziebland S. (2000) Analysing qualitative data. *British Medical Journal* **320**, 114–116.

Poses R. & Isen A. (1998) Qualitative research in medicine and health care: questions and controversy. *Journal of General Internal Medicine* **13**, 32–38.

Richardson L. (2000) Writing: a method of inquiry. In *Handbook of Qualitative Research*, 2nd edn (eds N.K. Denzin & Y.S. Lincoln), pp. 923–948. Sage Publications, Thousand Oaks, London, New Delhi.

Rowan M. & Huston P. (1997) Qualitative research articles: information for authors and peer reviewers. *Canadian Medical Association Journal* **157**, 1442–1446.

Seale C. (1999) *The Quality of Qualitative Research.* Sage Publications, Thousand Oaks, London, New Delhi.

Smith J. & Deemes D. (2000) The problem of criteria in the age of relativism. In *Handbook of Qualitative Research*, 2nd edn (eds N.K. Denzin & Y. Lincoln), pp. 877–896. Sage Publications, Thousand Oaks, London, New Delhi.

Sofaer S. (1999) Qualitative methods: what are they and why use them? *Health Services Research* **34** (Suppl. Part 2), 1101–1118.

Sparkes A. (2001) Myth 94: qualitative health researchers will agree about validity. *Qualitative Health Research* **11** (4), 538–552.

Spencer J. (1993) The usefulness of qualitative methods in rehabilitation: issues of meaning, of context and of change. *Archives of Physical and Medical Rehabilitation* **74**, 119–136.

Upshur R. E., VanDenKerkhof E.G. & Goel V. (2001) Meaning and measurement: an inclusive model of evidence in health care. *Journal of Evaluation in Clinical Practice* **7** (2), 91–96.

Ward M. (1993) Study design in qualitative research: a guide to assessing quality. *Journal of General Internal Medicine* **8**, 107–109.

Whittemore R., Chase S. & Mandle C. (2001) Validity in qualitative research. *Qualitative Health Research* **11** (4), 522–537.

Willms D.G., Best J.A., Taylor D.W., Gibert J.R., Wilson D.M.C., Lindsay E.A. & Singer J. (1997) A systematic approach for using qualitative methods in primary prevention research. *Medical Anthropology Quarterly* **4**, 391–409.

