



A Strategy to Identify Critical Appraisal Criteria for Primary Mixed-Method Studies

JOANNA E. M. SALE*

Health Research Methodology, Faculty of Health Sciences, McMaster University, 1280 Main St. W, Hamilton, ON L8N 1G8

KEVIN BRAZIL

Clinical Epidemiology and Biostatistics, Faculty of Health Sciences, McMaster University, 105 Main St. E., Level P1 Hamilton, ON L8N 1G8

Abstract. The practice of mixed-methods research has increased considerably over the last 10 years. While these studies have been criticized for violating quantitative and qualitative paradigmatic assumptions, the methodological quality of mixed-method studies has not been addressed. The purpose of this paper is to identify criteria to critically appraise the quality of mixed-method studies in the health literature. Criteria for critically appraising quantitative and qualitative studies were generated from a review of the literature. These criteria were organized according to a cross-paradigm framework. We recommend that these criteria be applied to a sample of mixed-method studies which are judged to be exemplary. With the consultation of critical appraisal experts and experienced qualitative, quantitative, and mixed-method researchers, further efforts are required to revise and prioritize the criteria according to importance.

Key words: mixed methods, multiple methods, qualitative methods, quantitative methods, critical appraisal

1. Introduction

The practice of mixed-methods research has increased considerably over the last 10 years, as seen in numerous articles, chapters, and books published (Caracelli and Greene, 1993; Carcelli and Riggin, 1994; Casebeer and Verhoef, 1997; Datta, 1997; Droitcour, 1997; Greene and Caracelli, 1997; House, 1994; Morgan, 1998; Tashakkori and Teddlie, 1998). Journal series have been devoted to this topic as well (see Volume 19 of *Health Education Quarterly* (1992); Number 74 of *New Directions for Evaluation* (1997); Volume 34 of *Health Services Research* (1999)). Despite being criticized for violating quantitative and qualitative paradigmatic assumptions, the methodological quality of mixed-method studies has not been examined. While one could argue that the quality of mixed-method studies cannot be assessed until the quantitative-qualitative philosophical debate is resolved, it will

* Author for correspondence: E-mail: joanna.sale@sw.ca

be demonstrated that under the assumption that methods are linked to paradigms, it is possible to identify criteria to critically appraise mixed-method studies.

There are difficulties with developing criteria to critically appraise mixed-method studies. The definitions for qualitative and quantitative methods as well as the paradigms linked with these methods vary considerably. The critical appraisal criteria proposed for the methods often overlap depending on the dominant paradigm view taken. Finally, the language of criteria is often vague, requiring the inexperienced reader to make judgement calls (Marshall, 1990). This paper will address the above difficulties. It will attempt to clarify the language applied to critical appraisal and the assumptions of the authors will be presented.

The purpose of this paper is to identify criteria to critically appraise the methods of primary studies in the health literature which employ mixed-methods. A mixed-methods study is defined as one in which quantitative and qualitative methods are combined in a single study. A primary study is defined as one that contains original data on a topic (Oxman et al., 1993). The methods refer to the procedures of the study. Before conducting the literature review, the authors will first outline their assumptions concerning methods linked to paradigms, criteria linked to paradigms, and the quantitative and qualitative method. Second, the concept of critical appraisal as it applies to quantitative and qualitative methods will be reviewed. Third, the theoretical framework upon which the authors rely will be described.

1.1. ASSUMPTIONS

1.1.1. *Methods Linked to Paradigms*

A paradigm is a patterned set of assumptions concerning reality (ontology), knowledge of that reality (epistemology), and the particular ways of knowing that reality (methodology) (Guba, 1990). The assumption of linking methods to a paradigm or paradigms is not a novel idea. It has been argued that different paradigms should be applied to different research questions within the same study and that the paradigms should remain clearly identified and separated from each other (Kuzel and Like, 1991). In qualitative studies, the methods are most often linked with the paradigms of interpretivism (Kuzel and Like, 1991, Altheide and Johnson, 1994; Secker et al., 1995), and constructivism (Guba and Lincoln, 1994a). However, other paradigms for qualitative methods have been proposed such as positivism (Devers, 1999), post-positivism (Marshall, 1990; Devers, 1999), postmodernism (Creswell, 1998), and critical theory (Creswell, 1998). Historically, quantitative methods have been linked with the paradigm of positivism. In recent years, quantitative methods have been viewed as evolving into the practice of post-positivism (Popper, 1959; Reichardt and Kallis, 1994). In relation to critical appraisal, we view positivism and post-positivism as compatible. For the remainder of the text, the term positivism will apply to both positivism and post-positivism.

This paper assumes that quantitative methods are linked with the paradigm of positivism and that qualitative methods are linked with the paradigms of construct-

ivism and interpretivism. This is an initial proposal for paradigms. We realize that critical appraisal criteria may change based on other paradigmatic assumptions.

1.1.2. *Criteria Linked to Paradigms*

Under the assumption that methods are linked to paradigms, it follows that the criteria to critically appraise these methods should also be linked to paradigms. The application of separate criteria for each method and paradigm is supported by many researchers (Lincoln and Guba, 1995, 1986; Burns, 1989; Forchuk and Roberts, 1993; Foster, 1997; Morse, 1991; Wilson, 1985; Yonge and Slewin, 1988). Method-specific criteria provides an alternative to blending methodological assumptions (Foster, 1997). Method-specific criteria also implies that each method is valued equally. While it is assumed that criteria are linked to paradigms, it is acknowledged that not all researchers (Altheide and Johnson, 1994, Guba and Lincoln, 1994a) maintain this perspective.

1.1.3. *The Quantitative Method*

Under the assumption that quantitative methods are based on the paradigm of positivism, the underlying view is that all phenomenon can be reduced to empirical indicators which represent the truth. The ontological position of this paradigm is that there is only one truth, an objective reality that exists independent of human perception. Epistemologically, the investigator and investigated are independent entities. Therefore, an investigator is capable of studying a phenomenon without influencing it or being influenced by it; "inquiry takes place as through a one way mirror" (Guba and Lincoln, 1994a).

1.1.4. *The Qualitative Method*

Under the assumption that qualitative methods are based on the paradigm of interpretivism and constructivism, multiple realities or multiple truths exist based on one's construction of reality. Reality is socially constructed (Berger and Luckman, 1966) and so is constantly changing. On an epistemological level, there is no access to reality independent of our minds, no external referent by which to compare claims of truth (Smith, 1983). The investigator and the object of study are interactively linked so that findings are mutually created within the context of the situation which shapes the inquiry (Guba and Lincoln, 1994b; Denzin and Lincoln, 1994). This suggests that reality has no existence prior to the activity of investigation, and reality ceases to exist when we lose interest in it (Smith, 1983). The emphasis of qualitative research is on process and meanings (Denzin and Lincoln, 1994).

1.2. THE CONCEPT OF CRITICAL APPRAISAL FOR QUANTITATIVE AND QUALITATIVE METHODS

Guidelines for critically appraising quantitative methods exist in many peer-reviewed journals. For example, BMJ has published criteria for the critical appraisal of manuscripts which are submitted to it. More recently, qualitative journals have been publishing their own set of criteria. For example, the Canadian Journal of Public Health now has published guidelines for qualitative research papers. However, it has been argued that the idea of static criteria may violate the assumptions of the paradigms to which qualitative methods belong (Smith, 1990; Garratt and Hodkinson, 1998; Sparkes, 2001). In this paper, we posit that there are static elements of qualitative methods which should be subject to criticism.

1.3. THEORETICAL FRAMEWORK – TRUSTWORTHINESS AND RIGOR PROPOSED BY LINCOLN AND GUBA (1985, 1986)

Our paper relies on Lincoln and Guba's (1985, 1986) framework of trustworthiness and rigor. This framework which was selected due to its cross-paradigm appeal, guided the organization of the criteria for critically appraising mixed-method studies generated by the literature review. Although Lincoln and Guba intended the framework to reflect conventional (positivistic) versus naturalistic paradigms, we feel the framework encompasses constructivism and interpretivism under that which is "naturalistic". (Guba and Lincoln themselves have used the term constructivism to refer to naturalistic inquiry (Guba and Lincoln, 1994c).) The term "trustworthiness", referred to as "goodness criteria" (Guba and Lincoln, 1994b), parallels the term "rigor" used in quantitative methods (Lincoln and Guba, 1986). Trustworthiness and rigor encompass four goals which apply to the respective methods and paradigms:

1. Truth value – internal validity for quantitative methods versus credibility for qualitative methods;
2. Applicability – external validity for quantitative methods versus transferability or fittingness for qualitative methods;
3. Consistency – reliability for quantitative methods versus dependability for qualitative methods;
4. Neutrality – objectivity for quantitative methods versus confirmability for qualitative methods.

Consistent with these goals, we propose to extend Lincoln and Guba's theoretical framework to include parallel quantitative and qualitative critical appraisal criteria to which mixed-method studies can be applied.

2. Method

The purpose of the literature review was to generate criteria which have been proposed for critically appraising the methods of quantitative, qualitative, and mixed-method studies. Although the purpose of this review was not to develop a measure of critical appraisal, the initial search for criteria was based on the principles of item generation (Berk, 1979) used in measurement theory. Following the principles of item generation, inclusion and exclusion rules were applied to the criteria considered.

2.1. SEARCH STRATEGY

Articles judged to be exemplary by the authors were extracted from personal files and the files of colleagues, and were reviewed for MESH headings and keywords. The Subject Headings Indices for Medline, Cinahl, and PsychINFO were searched for terms referring to quantitative methods, qualitative methods, mixed-methods research, and critical appraisal. Textwords (words that appear in the titles or abstracts) were specified where there were no subject headings for quantitative, qualitative, and mixed-method studies. Keywords/textwords were assigned to a particular block; blocks were constructed using the Boolean 'OR' operator. Critical appraisal terms were then combined with the blocks for qualitative, quantitative, or mixed-methods using the Boolean 'AND' operator. Table I outlines the search strategy which included all available years of the selected databases.

The databases Medline (1966-2000), HealthStar (1975-1999), CINAHL (1982-1999), PsycINFO (1967-2000/2001), Sociological Abstracts (1963-2000), Social Sciences Index (1983-1999), ERIC (1966-1999/09), and Current Contents (Jan 3, 1999-Feb 14, 2000) were searched. A "words anywhere" search was conducted in WebSPIRS as there are limited subject headings in the included databases. The terms for the critical appraisal block were limited in WebSPIRS as other terms such as "guidelines", "criteria", "assessment", and "quality" produced a high volume of hits which were not relevant. Articles were limited to those in the English language. To complement the literature search strategy, reference lists of articles were also perused for relevant articles and the journal *Qualitative Health Research* for the years 1997-2001 was hand-searched. A total of four articles could not be located and therefore were not retrieved.

2.2. INCLUSION AND EXCLUSION CRITERIA

The articles retrieved were reviewed and the criteria for critical appraisal were subjected to inclusion and exclusion criteria. Criteria appearing in a list or scale format were extracted from this format and screened independently for inclusion and exclusion. The criteria had to refer to the methods or methods sections of the research. Finally, the criteria had to be operational in that a reader would be able to

Table I. Search strategy for critically appraising mixed-method studies

Medline, HealthStar (limit to nonmedline) (MESH)	CINAHL	WebSPIRS (includes PsycINFO, Sociological Abstracts, Social Sciences Index)	ERIC	Current Contents
*Nursing Research	*Quality Assessment	critical appraisal		
*Research Design/st [Standards]	/mt [Methods] *Quality	standards		
*Health Services Research/mt [Methods]	Assessment/st [Standards] *Research/ev			
*Program Development/mt [Methods]	[Evaluation] *Research/st [Standards]			
*Program Evaluation/mt [Methods]	*Study Design			
*Research/mt [Methods]				
*Research Design/mt [Methods]				
AND				
quantitative.tw	*Qualitative studies	quantitative	quantitative [&]	quantitative
qualitative.tw	quantitative.tw	qualitative	qualitative	methods
quantitative	*Multimethod	quantitative	mixed [&] method	qualitative
qualitative.tw	Studies	qualitative	multimethod	methods
mixed method.tw	mixed method.tw	mixed method	multiple methods	mixed method
mixed methods.tw		mixed methods	mixed methodologies	multimethod
mixed		mixed	quantitative/qualitative	multiple methods
methodologies.tw		methodologies	qualitative/quantitative	
multimethod.tw.		multimethod	quantitative-qualitative	
multiple methods.tw		multiple methods	qualitative-quantitative	

* limit to focus.

review an article and determine whether criteria were met or not. (The authors operationalized the language typically applied in the critical appraisal literature; this document is available upon request). Criteria which required judgement calls on the part of the reader such as “is the purpose of the study one of discovery and description, conceptualization, illustration, or sensitization?” (Cobb and Hagemaster, 1987) and “was the design of the study sensible?” were excluded. Other exclusion criteria are noted in Table II. Where appropriate, examples of the exclusions are given.

Table II. Exclusion criteria

1.	<i>Criteria which do not fit the respective paradigm assumptions</i> as stated – in cases where assumptions were not stated, judgement calls were made based on the criteria themselves. For example, Morse (1991) suggests selecting a large enough random sample in qualitative research to compensate for the exclusion of those who are not good informants. Under the assumption of constructivism and interpretivism, random sampling is not employed.
2.	<i>Criteria which are disease-specific</i>
3.	<i>Criteria specific to design or tradition</i> within each method and not applicable to other designs or traditions within that method – for example, criteria specific to randomized controlled trials (RCTs), case studies, or ethnographies
4.	<i>Criteria specific to review articles</i>
5.	<i>Criteria specific to statistical tests</i>
6.	<i>Criteria concerning the credibility of the researcher</i>
7.	<i>Criteria specific to the quality and reporting</i> of the literature review
8.	<i>Criteria for the elements of a proposal</i>
9.	<i>Criteria specific to the following topics in medicine:</i> etiology, improving clinical performance, diagnosis, prognosis, therapy, causation, economics of health care programs or interventions, practice guidelines
10.	<i>Criteria specific to the importance, relevance, or significance</i> of the research
11.	<i>Criteria specific to measurement scales</i>

3. Results

As we suspected, no criteria for critically appraising mixed-method studies were found. Table III is a review of the criteria for critically appraising quantitative and qualitative methods separately. Because this paper assumes that methods and criteria are linked to paradigms, the criteria appear separately. As we have proposed, the organization of the table is based on the framework of trustworthiness and rigor proposed by Lincoln and Guba (1985, 1986). The criteria have been grouped accordingly but it is acknowledged that there may be some overlap of criteria across the four goals.

4. Discussion

This paper has attempted to propose criteria for the purpose of critically appraising primary mixed-method studies. It is acknowledged that this is not a simple task; rather than a final product, this paper is an initial step in the process of identifying such criteria. It is necessary that the “final product” is manageable and realistic to achieve in the length of a peer-reviewed journal article. We therefore envision the final product to be a reduced version of Table III combined with criteria specific to mixed-methods such as acknowledgement of paradigm assumptions/differences and identification of the mixed-methods design.

Table III. Review of critical appraisal criteria for quantitative and qualitative methods

Goals of criteria	Qualitative methods	Quantitative methods
Truth value (Credibility vs. Internal Validity)	<p>Triangulation of sources (Lincoln and Guba, 1985; Patton, 1990, 1999; Yin, 1999)</p> <p>Triangulation of methods (Lincoln and Guba, 1985; Patton, 1990, 1999; Elliot et al., 1999)</p> <p>Triangulation of investigators (Lincoln and Guba, 1985; Patton, 1990, 1999)</p> <p>Triangulation of theory/perspective (Patton, 1990, 1999)</p> <p>Peer debriefing (Lincoln and Guba, 1985; Inui and Frankel, 1991)</p> <p>Negative case analysis or searching for disconfirming evidence (Marshall, 1990; Lincoln and Guba, 1985; Patton, 1990, 1999; Reid, 1996)</p> <p>Member checks (Lincoln, and Guba, 1985; Morse, 1991; Elliott et al., 1999; Inui and Frankel, 1991; Reid, 1996)</p> <p>Use of quotations (Secker et al., 1995)</p> <p>Informed consent stated (Lincoln and Guba, 1985; Burns, 1989; Elliott et al., 1999)</p> <p>Ethical review or human subject review undertaken (Morse, 1991; Elliott et al., 1999)</p> <p>Statement that confidentiality protected (Forchuk and Roberts, 1993; Elliott et al., 1999)</p> <p>Consent procedures described (Forchuk and Roberts, 1993; Morse, 1991)</p>	<p>Extraneous (Haughey, 1994a) or confounding variables (Fowkes and Fulton, 1991) identified</p> <p>Extraneous or confounding variable(s) or baseline differences controlled for in the analysis (Greenhalgh, 1997a; Haughey, 1994a; Fowkes and Fulton, 1991; Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981b; Laupacis et al., 1994)</p> <p>Statement about comparability of control group to intervention group at baseline (Guyatt et al., 1994; Levine et al., 1994)</p> <p>Statement that comparison group treated equally to aside from intervention (Guyatt et al., 1994)</p> <p>Informed consent stated (Elliott et al., 1999; Haughey, 1994b)</p> <p>Ethical review undertaken (Elliott et al., 1999; Haughey, 1994b)</p> <p>Statement that confidentiality protected (Elliott et al., 1999; Haughey, 1994b)</p>
Applicability (Transferability/ Fittingness vs. External Validity/ Generalizability)	<p>Statement of purpose (Burns, 1989; Cobb and Hagemaster, 1987; Elliott et al., 1999)</p> <p>Statement of research question(s) (Marshall, 1990; Forchuk and Roberts, 1993; Morse, 1991; Cobb and Hagemaster, 1987; Morse, 1991, Inui and Frankel, 1991; Reid, 1996)</p> <p>Phenomenon of study stated (Burns, 1989; Sale et al., 2002)</p> <p>Rationale for the use of qualitative methods (Marshall, 1990)</p>	<p>Statement of purpose (Elliott et al., 1999)</p> <p>Objective of study explicitly stated or described (Fowkes and Fulton, 1991; Gardner et al., 1986)</p> <p>Description of intervention if appropriate (Haughey, 1994a)</p> <p>Outcome measure(s) defined (Greenhalgh, 1997; Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981b; Laupacis et al., 1994)</p>

Table III. Continued

Goals of criteria	Qualitative methods	Quantitative methods
	Rationale for the tradition within qualitative methods (Forchuk and Roberts, 1993)	Assessment of outcome blinded (Greenhalgh, 1997a; Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981b)
	Description of study context (Forchuk and Roberts, 1993; Reid, 1996) or setting (Lincoln and Guba, 1985)	Description of setting or conditions under which data collected (Haughey, 1994a, b, Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981a)
	Statement of how setting was selected (Greenhalgh and Taylor, 1997)	Design stated explicitly i.e. case study, cross-sectional study, cohort study, RCT (Haughey, 1994a; Fowkes and Fulton, 1991)
	Sampling procedure described (Burns, 1989; Popay et al., 1998; Greenhalgh and Taylor, 1997)	Subject recruitment or sampling selection described (Greenhalgh, 1997a; Laupacis et al., 1994; Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981a)
	Justification or rationale for sampling strategy (Inui and Frankel, 1991; Reid, 1996)	Sample randomly selected (Greenhalgh and Taylor, 1997; Morse, 1986)
	Description of participants or informants (Burns, 1989; Forchuk and Roberts, 1993; Elliott et al., 1999)	Inclusion and exclusion criteria for subject selection stated explicitly (Greenhalgh, 1997a; Haughey, 1994a; Fowkes and Fulton, 1991; Dunn, 1991)
	Data gathering procedures described (Lincoln and Guba, 1985; Forchuk and Roberts, 1993; Elliott et al., 1999)	Study population defined or described (Greenhalgh, 1997a; Haughey, 1994a; Laupacis et al., 1994; Dunn, 1991)
	Audiotaping procedures described (Reid, 1996)	Source of subjects stated i.e. sampling frame identified (Morse, 1991; Gardner et al., 1986)
	Transcription procedures described (Reid, 1996)	Source of controls stated (Fowkes and Fulton, 1991)
	Field note procedures described (Reid, 1996)	Selection of controls described (Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981a)
	Data analysis described (Lincoln and Guba, 1985; Forchuk and Roberts, 1993)	Control or comparison group described (Fowkes and Fulton, 1991)
	Coding techniques described (Lincoln and Guba, 1985; Burns, 1989; Morse, 1991; Reid, 1996)	Statement about nonrespondents (Fowkes and Fulton, 1991; or dropouts or deaths (Greenhalgh, 1997a; Fowkes and Fulton, 1991)
	Data collection to saturation specified (Morse, 1991)	
	Statement that reflexive journals or logbooks kept (Inui and Frankel, 1991; Koch, 1994)	
	Description of raw data (Lincoln and Guba, 1985)	

Table III. Continued

Goals of criteria	Qualitative methods	Quantitative methods
Consistency (Dependability vs. Reliability) Neutrality (Confirmability vs. Objectivity)	External audit of process (Lincoln and Guba, 1985) External audit of data and reconstructions of the data (Lincoln and Guba, 1985) Bracketing (Secker et al., 1995; Burns, 1989; Patton, 1999) Statement of researcher's assumptions (Marshall, 1990; Elliott et al., 1999) or statement of researcher's perspective (Greenhalgh and Taylor, 1997)	Missing data addressed (Fowkes and Fulton, 1991) Power calculation to assess adequacy of sample size (Fowkes and Fulton, 1991; Gardner et al., 1986) or sample size calculated for adequate power (Greenhalgh, 1997a) Statistical procedures referenced or described (Gardner et al., 1986) p values stated (Greenhalgh, 1997b) Confidence intervals given for main results (Laupacis et al., 1994; Guyatt et al., 1994; Levine et al., 1994; Gardner et al., 1986; Greenhalgh, 1997b) Data gathering procedures described (Elliott et al., 1999) Data collection instruments or source of data described (Greenhalgh, 1997a; Haughey, 1994a) At least one hypothesis stated (Dunn, 1991) Both statistical and clinical significance acknowledged (Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre, 1981c)
		Standardization of observers described (Dunn, 1991)

An alternative to this recommendation might be the criteria shared by both methods such as the description of participants/study population and the description of the study context or setting. However, it is the authors' opinion that the methodological evaluation of mixed-method studies should not be limited to those criteria which are shared. Paradigmatic differences between quantitative and qualitative methods imply that criteria shared by both methods may not be equally important to each method.

It is interesting to note that most of the criteria generated by the literature review fit into the first 2 categories of Lincoln and Guba's framework. That is, most of these criteria address the goals of truth value (credibility vs. internal validity) and applicability (transferability vs. external validity). It is possible that the remaining two goals are not as relevant for critical appraisal or that it is more difficult to operationalize criteria which should be included under these goals. It is also possible that the criteria overlap and that those which are suitable for these goals are more suitable under one of the first 2 goals.

While this paper has attempted to identify criteria which are generic to mixed-method studies, future guidelines for criteria might be more specific depending upon the mixed-methods design. As mixed-methods research becomes more sophisticated, there may be varying combinations of criteria for the specific designs. For example, the criteria to critically appraise an ethnographic-randomized controlled trial design may differ from those for a phenomenological-case-control design.

The overall goal of proposing criteria for critically appraising mixed-method studies is to promote standards for guiding and assessing the methodological quality of studies of this nature. It is anticipated that, as with quantitative and qualitative research, there will eventually be guidelines for reporting and writing about mixed-method studies. It is hoped that this paper will stimulate future dialogue on the critical appraisal of such studies.

5. Future Considerations in the Pursuit of Critical Appraisal for Mixed-Method Studies

The application of criteria to mixed-method studies will rely upon reported criteria. It is possible that criteria will have been met but not reported due to editorial revisions or journal guidelines concerning the length of articles. However, one purpose of this paper is to improve the reporting of methods in mixed-method studies; it can be argued, therefore, that is appropriate to make recommendations based on criteria which have been reported.

This paper is limited by the judgements made by the authors concerning exclusion and inclusion of criteria. We believed that the proposed criteria for each method should be generic and applicable to other studies within that method. We also believed that criteria had to be operationally defined. Having said this, "description of study population" does not specify what elements of a study population should be described. For the purpose of this paper, any description of the sample

would allow this criteria to be met. All exclusions have been identified in Table I; it is possible that these exclusions may be challenged during the revision of criteria. The authors welcome such debate. In the spirit of furthering this debate, we recommend the following considerations:

The criteria in Table III should be further refined and then applied to a sample of mixed-method studies in the health literature. It is also recommended that experts in mixed-methods research be asked to identify mixed-method studies which they judge to be exemplary so that the criteria met by these studies can be assessed as well.

It is noted that there are few subject headings in the selected databases which reflect the key concepts of this paper. An attempt was made to devise a comprehensive search strategy which often relied on the use of textwords. The inclusiveness of this strategy is unknown. The search strategy also focused on the peer-reviewed literature; relevant books, conference proceedings, and unpublished manuscripts may have been missed.

This paper assumed that all criteria generated by the literature were equally important. However, it is possible that the criteria met by mixed-method studies might be those which are easiest to meet rather than those which are important. The revision of criteria should involve consultation with critical appraisal experts and experienced qualitative, quantitative, and mixed-method researchers. It is proposed that this panel of experts and experienced researchers rate the criteria according to importance and that this rating be taken into account during the revision process.

References

- Altheide, D. L. & Johnson, J. M. (1994). Criteria for assessing interpretive validity in qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, pp. 485–499.
- Berger, P. L. & Luckmann, T. (1996). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Doubleday.
- Berk, R. A. (1979). The construction of rating instruments for faculty evaluation: A review of the methodological issues. *J. High Educ.* 50(5): 650–670.
- Burns, N. (1989). Standards for qualitative research. *Nurs. Sci. Q.* 2(1): 44–52.
- Caracelli, V. J. & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educ. Eval. Policy. An.* 15: 195–207.
- Caracelli, V. J. & Riggan, L. J. C. (1994). Mixed method evaluation: Developing quality criteria through concept mapping. *Eval. Pract.* 15: 139–152.
- Casebeer, A. L. & Verhoef, M. J. (1997). Combining qualitative and quantitative research methods: Considering the possibilities for enhancing the study of chronic diseases. *Chronic Dis. Can.* 18: 130–135.
- Cobb, A. K. & Hagemaster, J. N. (1987). Ten criteria for evaluating qualitative research proposals. *J. Nurs. Educ.* 26(4): 138–143.
- Creswell, J. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Thousand Oaks, CA: Sage Publications.
- Datta, L. (1997). Multimethod evaluations: Using case studies together with other methods. In: E. Chelimsky & W. R. Shadish (eds), *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage Publications, pp. 344–359.

- Denzin, N. K. & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, pp. 1–17.
- Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre (1981a). How to read clinical journals: II. To learn about a diagnostic test. *CMAJ* 124: 703–710.
- Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre (1981b). How to read clinical journals: III To learn the clinical course and prognosis of disease. *CMAJ* 124: 869–872.
- Department of Clinical Epidemiology and Biostatistics, McMaster University Health Sciences Centre (1981c). How to read clinical journals: V: To distinguish useful from useless or even harmful therapy. *CMAJ* 124: 1156–1162.
- Devers, K. J. (1999). How will we know ‘good’ qualitative research when we see it? Beginning the dialogue in health services research. *Health Serv. Res.* 45(5 Part II): 1153–1188.
- Droitcour, J. A. (1997). Cross design synthesis: Concept and application. In: E. Chelimsky & W. R. Shadish (eds), *Evaluation for the 21st Century: A Handbook*. Thousand Oaks, CA: Sage Publications, pp. 360–372.
- Dunn, E. V. (1991). Basic standards for analytic studies in primary care research. In: P. G. Norton, M. Stewart, F. Tudiver, M. J. Bass & E. V. Dunn (eds), *Primary Care Research*. Newbury Park, CA: Sage Publications, pp. 78–96.
- Elliott, R., Fischer, C. T. & Rennie, D. L. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *Br. J. Clin. Psychol.* 38(3): 215–229.
- Ford-Gilboe, M., Campbell, J. & Berman, H. (1995). Stories and numbers: Coexistence without compromise. *Adv. Nurs. Sci.* 13(1): 14–26.
- Forchuk, C. & Roberts, J. (1993). How to critique qualitative research articles. *Can. J. Nurs. Res.* 25(4): 47–56.
- Foster, R. L. (1997). Addressing epistemological and practical issues in multimethod research: A procedure for conceptual triangulation. *Adv. Nurs. Sci.* 20(2): 1–12.
- Fowkes, F. G. R. & Fulton, P. M. (1991). Critical appraisal of published research: Introductory guidelines. *BMJ* 302: 1136–1140.
- Gardner, M. J., Machin, D. & Campbell, M. J. (1986). Use of check lists in assessing the statistical content of medical studies. *BMJ* 292: 810–812.
- Garratt, D. & Hodkinson, P. (1998). Can there be criteria for selecting research criteria? – A hermeneutical analysis of an inescapable dilemma. *Qualitative Inquiry* 4(4): 515–539.
- Greene, J. C. & Caracelli, V. J. (eds) (1997). Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. *New Directions for Program Evaluation*. San Francisco: Jossey-Boss Publishers.
- Greenhalgh, T. (1997a). Assessing the methodological quality of published papers. *BMJ* 315(7103): 305–308.
- Greenhalgh, T. (1997b). Statistics for the nonstatistician. II: ‘Significant’ relations and their pitfalls. *BMJ* 315(7105): 422–425.
- Greenhalgh, T. & Taylor, R. (1997). Papers that go beyond numbers (qualitative research). *BMJ* 315: 740–743.
- Guba, E. G. (1990). The alternative paradigm dialog. In: E. G. Guba (ed.), *The Paradigm Dialog*. Newbury Park, CA: Sage Publications, pp. 17–30.
- Guba, E. G. & Lincoln, Y. S. (1994a). Competing paradigms in qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, p. 110.
- Guba, E. G. & Lincoln, Y. S. (1994b). Competing paradigms in qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications, pp. 105–117.

- Guba, E. G. & Lincoln, Y. S. (1994c). Competing paradigms in qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage Publications.
- Guyatt, G. H., Sackett, D. L., Cook, D.J., for the Evidence-Based Working Group. (1994). Users' guide to the medical literature. II: How to use an article about therapy or prevention, B. What were the results and will they help me in caring for my patients? *JAMA* 271(1): 59–63.
- Haughey, B. P. (1994a). Evaluating quantitative research designs: Part 1. *Crit. Care Nurse* 14(5): 100–102.
- Haughey, B. P. (1994b). Evaluating quantitative research designs: Part 2. *Crit. Care Nurse* 14(6): 69–72.
- Health Services Research (1999). 34(5).
- Health Educational Quarterly (1992). 19(1).
- House, E. R. (1994). Integrating the qualitative and quantitative. In: C. S. Reichardt & S. F. Rallis (eds), *The Qualitative-Quantitative Debate: New Perspectives*. San Francisco: Jossey-Boss Publishers.
- Inui, T. S. & Frankel, R. M. (1991). Evaluating the quality of qualitative research. *J. Gen. Internal Med.* 6: 485–486.
- Koch, T. (1994). Establishing rigour in qualitative research: The decision trail. *J. Adv. Nurs.* 19: 976–986.
- Kuzel, A. J. & Like, R. C. (1991). Standards of trustworthiness for qualitative studies in primary care. In: P. G. Norton, M. Steward, F. Tudiver, M. J. Bass & E. V. Dunn (eds), *Primary Care Research*. Newbury Park, CA: Sage Publications; pp. 138–158.
- Laupacis, A., Wells, G., Richardson, S. & Tugwell, P., for the Evidence-Based Medicine Working Group (1994). Users' guides to the medical literature: V. How to use an article about prognosis. *JAMA* 272(3): 234–237.
- Levine, M., Walter, S., Lee, H., Haines, T., Holbrook, A., Moyer, V., for the Evidence-Based Working Group. (1994). Users' guide to the medical literature. IV. How to use an article about harm. *JAMA* 271(20): 1615–1619.
- Lincoln, Y.S. & E. G. Guba. (1985). *Naturalistic Inquiry*. Newbury Park, CA: Sage Publications.
- Lincoln, Y. S. & Guba, E. G. (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. In: D. D. Williams (ed.), *Naturalistic Evaluation. New Directions for Program Evaluation, No. 90*. San Francisco: Jossey-Boss, pp. 78–84.
- Marshall C. (1990). Goodness criteria: Are they objective or judgement calls? In: E. G. Guba (ed.), *The Paradigm Dialog*. Newbury Park, CA: Sage Publications, pp. 188–197.
- Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qual. Health Res.* 8: 362–376.
- Morse, J. M. (1991). Evaluating qualitative research. *Qual. Health Res.* 1(3): 283–286.
- Morse, J. M. (1986). Quantitative and qualitative research: Issues in sampling. In: P. L. Chin (ed.), *Nursing Research Methodology: Issues and Implementation*. Rockville, MA: Aspen Publishers.
- New Directions for Evaluation: Advances in Mixed-Method Evaluation: The Challenges and Benefits of Integrating Diverse Paradigms 1997; p. 74.
- Oxman, A.D., Sackett, D. L. & Guyatt, G. H. (1993). For the Evidence-Based Medicine Working Group. Users' guides to the medical literature: 1. How to get started. *JAMA* 270(17): 2093–2095.
- Patton, M. Q. (1990). Enhancing the quality and credibility of qualitative analysis. In: *Qualitative Evaluation and Research Methods*, 2nd edn. Newbury Park, CA: Sage Publications.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Serv. Res.* 34(5 Part II): 1189–1208.
- Popay, J., Rogers, A. & Williams, G. (1998). Rationale and standards for the systematic review of qualitative literature in health services research. *Qual. Health Res.* 8(3): 341–351.
- Popper, K. R. (1959). *The Logic of Scientific Discovery*. New York: Basic Books.

- Reichardt, C. S. & Rallis, S.F. (1994). Qualitative and quantitative inquiries are not incompatible: A call for a new partnership. In: C. S. Reichardt & S. F. Rallis (eds), *The Qualitative-Quantitative Debate: New Perspectives*. San Francisco: Jossey-Boss Publishers, pp. 85–92.
- Reid, A. (1996). What we want: Qualitative research. *Can. Fam. Physician* 42: 387–389.
- Sale, J. E. M., Lohfeld, L. & Brazil, K. (2002). Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality & Quantity* 36: 43–53.
- Secker, J., Wimbush, E., Watson, J. & Milburn, K. (1995). Qualitative methods in health promotion research: Some criteria for quality. *Health Educ. J.* 54: 74–87.
- Smith, J. K. (1983). Quantitative versus qualitative research: An attempt to clarify the issue. *Educational Researcher* 12: 6–13.
- Smith, J. K. (1990). Alternative research paradigms and the problem of criteria. In: E. G. Guba (ed.), *The Paradigm Dialog*. Newbury Park, CA: Sage Publications, pp. 167–187.
- Sparkes, A. C. (2001). Myth 94: Qualitative health researchers will agree about validity. *Qual. Health Res.* 11(4): 538–552.
- Tashakkori, A. & Teddlie, C. (1998). *Mixed Methodology: Combining Qualitative and Quantitative Approaches*. Thousand Oaks, CA: Sage Publications.
- Wilson, H. S. (1985). Qualitative studies: From observations to explanations. *Journal Nurs. Adm.* 8–10.
- Yin, R. K. (1999). Enhancing the quality of case studies in health services research. *Health Serv. Res.* 34(5 Part II): 1209–1224.
- Yonge, O. & Slewin, L. (1988). Reliability and validity: Misnomers for qualitative research. *Can. J Nurs. Res.* 20(2): 61–67.

