

Validity, reliability, and generalizability in qualitative research

Lawrence Leung^{1,2}

¹Department of Family Medicine and ²Centre of Studies in Primary Care, Queen's University, Kingston, Ontario, Canada

ABSTRACT

In general practice, qualitative research contributes as significantly as quantitative research, in particular regarding psycho-social aspects of patient-care, health services provision, policy setting, and health administrations. In contrast to quantitative research, qualitative research as a whole has been constantly critiqued, if not disparaged, by the lack of consensus for assessing its quality and robustness. This article illustrates with five published studies how qualitative research can impact and reshape the discipline of primary care, spiraling out from clinic-based health screening to community-based disease monitoring, evaluation of out-of-hours triage services to provincial psychiatric care pathways model and finally, national legislation of core measures for children's healthcare insurance. Fundamental concepts of validity, reliability, and generalizability as applicable to qualitative research are then addressed with an update on the current views and controversies.

Keywords: Controversies, generalizability, primary care research, qualitative research, reliability, validity

Nature of Qualitative Research versus Quantitative Research

The essence of qualitative research is to make sense of and recognize patterns among words in order to build up a meaningful picture without compromising its richness and dimensionality. Like quantitative research, the qualitative research aims to seek answers for questions of "how, where, when who and why" with a perspective to build a theory or refute an existing theory. Unlike quantitative research which deals primarily with numerical data and their statistical interpretations under a reductionist, logical and strictly objective paradigm, qualitative research handles nonnumerical information and their phenomenological interpretation, which inextricably tie in with human senses and subjectivity. While human emotions and perspectives from both subjects and researchers are considered undesirable biases confounding results in quantitative research, the same elements are considered essential and inevitable, if not treasurable, in qualitative research as they invariably add extra dimensions and colors to enrich the corpus of findings. However, the issue of subjectivity and contextual ramifications has fueled incessant

controversies regarding yardsticks for quality and trustworthiness of qualitative research results for healthcare.

Impact of Qualitative Research upon Primary Care

In many ways, qualitative research contributes significantly, if not more so than quantitative research, to the field of primary care at various levels. Five qualitative studies are chosen to illustrate how various methodologies of qualitative research helped in advancing primary healthcare, from novel monitoring of chronic obstructive pulmonary disease (COPD) via mobile-health technology,^[1] informed decision for colorectal cancer screening,^[2] triaging out-of-hours GP services,^[3] evaluating care pathways for community psychiatry^[4] and finally prioritization of healthcare initiatives for legislation purposes at national levels.^[5] With the recent advances of information technology and mobile connecting device, self-monitoring and management of chronic diseases via tele-health technology may seem beneficial to both the patient and healthcare provider. Recruiting COPD patients who were given tele-health devices that monitored lung functions, Williams *et al.*^[1] conducted phone interviews and analyzed their transcripts via a grounded theory approach, identified themes

Access this article online

Quick Response Code:



Website:
www.jfmipc.com

DOI:
10.4103/2249-4863.161306

Address for correspondence: Prof. Lawrence Leung,
Centre of Studies in Primary Care, Queen's University,
220 Bagot Street, Kingston,
ON K7L 5E9, Canada.
E-mail: leungl@queensu.ca

which enabled them to conclude that such mobile-health setup and application helped to engage patients with better adherence to treatment and overall improvement in mood. Such positive findings were in contrast to previous studies, which opined that elderly patients were often challenged by operating computer tablets,^[6] or, conversing with the tele-health software.^[7] To explore the content of recommendations for colorectal cancer screening given out by family physicians, Wackerbarth, *et al.*^[2] conducted semi-structure interviews with subsequent content analysis and found that most physicians delivered information to enrich patient knowledge with little regard to patients' true understanding, ideas, and preferences in the matter. These findings suggested room for improvement for family physicians to better engage their patients in recommending preventative care. Faced with various models of out-of-hours triage services for GP consultations, Egbunike *et al.*^[3] conducted thematic analysis on semi-structured telephone interviews with patients and doctors in various urban, rural and mixed settings. They found that the efficiency of triage services remained a prime concern from both users and providers, among issues of access to doctors and unfulfilled/mismatched expectations from users, which could arouse dissatisfaction and legal implications. In UK, a care pathways model for community psychiatry had been introduced but its benefits were unclear. Khandaker *et al.*^[4] hence conducted a qualitative study using semi-structure interviews with medical staff and other stakeholders; adopting a grounded-theory approach, major themes emerged which included improved equality of access, more focused logistics, increased work throughput and better accountability for community psychiatry provided under the care pathway model. Finally, at the US national level, Mangione-Smith *et al.*^[5] employed a modified Delphi method to gather consensus from a panel of nominators which were recognized experts and stakeholders in their disciplines, and identified a core set of quality measures for children's healthcare under the Medicaid and Children's Health Insurance Program. These core measures were made transparent for public opinion and later passed on for full legislation, hence illustrating the impact of qualitative research upon social welfare and policy improvement.

Overall Criteria for Quality in Qualitative Research

Given the diverse genera and forms of qualitative research, there is no consensus for assessing any piece of qualitative research work. Various approaches have been suggested, the two leading schools of thoughts being the school of Dixon-Woods *et al.*^[8] which emphasizes on methodology, and that of Lincoln *et al.*^[9] which stresses the rigor of interpretation of results. By identifying commonalities of qualitative research, Dixon-Woods produced a checklist of questions for assessing clarity and appropriateness of the research question; the description and appropriateness for sampling, data collection and data analysis; levels of support and evidence for claims; coherence between data, interpretation and conclusions, and finally level of contribution of the paper. These criteria foster the 10 questions for the Critical Appraisal

Skills Program checklist for qualitative studies.^[10] However, these methodology-weighted criteria may not do justice to qualitative studies that differ in epistemological and philosophical paradigms,^[11,12] one classic example will be positivistic versus interpretivistic.^[13] Equally, without a robust methodological layout, rigorous interpretation of results advocated by Lincoln *et al.*^[9] will not be good either. Meyrick^[14] argued from a different angle and proposed fulfillment of the dual core criteria of "transparency" and "systematicity" for good quality qualitative research. In brief, every step of the research logistics (from theory formation, design of study, sampling, data acquisition and analysis to results and conclusions) has to be validated if it is transparent or systematic enough. In this manner, both the research process and results can be assured of high rigor and robustness.^[14] Finally, Kitto *et al.*^[15] epitomized six criteria for assessing overall quality of qualitative research: (i) Clarification and justification, (ii) procedural rigor, (iii) sample representativeness, (iv) interpretative rigor, (v) reflexive and evaluative rigor and (vi) transferability/generalizability, which also double as evaluative landmarks for manuscript review to the Medical Journal of Australia. Same for quantitative research, quality for qualitative research can be assessed in terms of validity, reliability, and generalizability.

Validity

Validity in qualitative research means "appropriateness" of the tools, processes, and data. Whether the research question is valid for the desired outcome, the choice of methodology is appropriate for answering the research question, the design is valid for the methodology, the sampling and data analysis is appropriate, and finally the results and conclusions are valid for the sample and context. In assessing validity of qualitative research, the challenge can start from the ontology and epistemology of the issue being studied, e.g. the concept of "individual" is seen differently between humanistic and positive psychologists due to differing philosophical perspectives:^[16] Where humanistic psychologists believe "individual" is a product of existential awareness and social interaction, positive psychologists think the "individual" exists side-by-side with formation of any human being. Set off in different pathways, qualitative research regarding the individual's wellbeing will be concluded with varying validity. Choice of methodology must enable detection of findings/phenomena in the appropriate context for it to be valid, with due regard to culturally and contextually variable. For sampling, procedures and methods must be appropriate for the research paradigm and be distinctive between systematic,^[17] purposeful^[18] or theoretical (adaptive) sampling^[19,20] where the systematic sampling has no *a priori* theory, purposeful sampling often has a certain aim or framework and theoretical sampling is molded by the ongoing process of data collection and theory in evolution. For data extraction and analysis, several methods were adopted to enhance validity, including 1st tier triangulation (of researchers) and 2nd tier triangulation (of resources and theories),^[17,21] well-documented audit trail of materials and processes,^[22-24] multidimensional analysis as concept- or case-orientated^[25,26] and respondent verification.^[21,27]

Reliability

In quantitative research, reliability refers to exact replicability of the processes and the results. In qualitative research with diverse paradigms, such definition of reliability is challenging and epistemologically counter-intuitive. Hence, the essence of reliability for qualitative research lies with consistency.^[24,28] A margin of variability for results is tolerated in qualitative research provided the methodology and epistemological logistics consistently yield data that are ontologically similar but may differ in richness and ambience within similar dimensions. Silverman^[29] proposed five approaches in enhancing the reliability of process and results: Refutational analysis, constant data comparison, comprehensive data use, inclusive of the deviant case and use of tables. As data were extracted from the original sources, researchers must verify their accuracy in terms of form and context with constant comparison,^[27] either alone or with peers (a form of triangulation).^[30] The scope and analysis of data included should be as comprehensive and inclusive with reference to quantitative aspects if possible.^[30] Adopting the Popperian dictum of falsifiability as essence of truth and science, attempted to refute the qualitative data and analytes should be performed to assess reliability.^[31]

Generalizability

Most qualitative research studies, if not all, are meant to study a specific issue or phenomenon in a certain population or ethnic group, of a focused locality in a particular context, hence generalizability of qualitative research findings is usually not an expected attribute. However, with rising trend of knowledge synthesis from qualitative research via meta-synthesis, meta-narrative or meta-ethnography, evaluation of generalizability becomes pertinent. A pragmatic approach to assessing generalizability for qualitative studies is to adopt same criteria for validity: That is, use of systematic sampling, triangulation and constant comparison, proper audit and documentation, and multi-dimensional theory.^[17] However, some researchers espouse the approach of analytical generalization^[32] where one judges the extent to which the findings in one study can be generalized to another under similar theoretical, and the proximal similarity model, where generalizability of one study to another is judged by similarities between the time, place, people and other social contexts.^[33] Thus said, Zimmer^[34] questioned the suitability of meta-synthesis in view of the basic tenets of grounded theory,^[35] phenomenology^[36] and ethnography.^[37] He concluded that any valid meta-synthesis must retain the other two goals of theory development and higher-level abstraction while in search of generalizability, and must be executed as a third level interpretation using Gadamer's concepts of the hermeneutic circle,^[38,39] dialogic process^[38] and fusion of horizons.^[39] Finally, Toye *et al.*^[40] reported the practicality of using "conceptual clarity" and "interpretative rigor" as intuitive criteria for assessing quality in meta-ethnography, which somehow echoed Rolfe's controversial aesthetic theory of research reports.^[41]

Food for Thought

Despite various measures to enhance or ensure quality of qualitative studies, some researchers opined from a purist ontological and epistemological angle that qualitative research is not a unified, but ipso facto diverse field,^[8] hence any attempt to synthesize or appraise different studies under one system is impossible and conceptually wrong. Barbour argued from a philosophical angle that these special measures or "technical fixes" (like purposive sampling, multiple-coding, triangulation, and respondent validation) can never confer the rigor as conceived.^[11] In extremis, Rolfe *et al.* opined from the field of nursing research, that any set of formal criteria used to judge the quality of qualitative research are futile and without validity, and suggested that any qualitative report should be judged by the form it is written (aesthetic) and not by the contents (epistemic).^[41] Rolfe's novel view is rebutted by Porter,^[42] who argued via logical premises that two of Rolfe's fundamental statements were flawed: (i) "The content of research report is determined by their forms" may not be a fact, and (ii) that research appraisal being "subject to individual judgment based on insight and experience" will mean those without sufficient experience of performing research will be unable to judge adequately – hence an elitist's principle. From a realism standpoint, Porter then proposes multiple and open approaches for validity in qualitative research that incorporate parallel perspectives^[43,44] and diversification of meanings.^[44] Any work of qualitative research, when read by the readers, is always a two-way interactive process, such that validity and quality has to be judged by the receiving end too and not by the researcher end alone.

In summary, the three gold criteria of validity, reliability and generalizability apply in principle to assess quality for both quantitative and qualitative research, what differs will be the nature and type of processes that ontologically and epistemologically distinguish between the two.

References

1. Williams V, Price J, Hardinge M, Tarassenko L, Farmer A. Using a mobile health application to support self-management in COPD: A qualitative study. *Br J Gen Pract* 2014;64:e392-400.
2. Wackerbarth SB, Tarasenko YN, Joyce JM, Haist SA. Physician colorectal cancer screening recommendations: An examination based on informed decision making. *Patient Educ Couns* 2007;66:43-50.
3. Egbunike JN, Shaw C, Porter A, Button LA, Kinnersley P, Hood K, *et al.* Streamline triage and manage user expectations: Lessons from a qualitative study of GP out-of-hours services. *Br J Gen Pract* 2010;60:e83-97.
4. Khandaker GM, Gandamaneni PK, Dibben CR, Cherukuru S, Cairns P, Ray MK. Evaluating care pathways for community psychiatry in England: A qualitative study. *J Eval Clin Pract* 2013;19:298-303.
5. Mangione-Smith R, Schiff J, Dougherty D. Identifying children's health care quality measures for Medicaid and CHIP: An evidence-informed, publicly transparent expert process. *Acad Pediatr* 2011;11:S11-21.

6. Hess R, Santucci A, McTigue K, Fischer G, Kapoor W. Patient difficulty using tablet computers to screen in primary care. *J Gen Intern Med* 2008;23:476-80.
7. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, *et al*. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: A qualitative study. *BMC Health Serv Res* 2012;12:220.
8. Dixon-Woods M, Shaw RL, Agarwal S, Smith JA. The problem of appraising qualitative research. *Qual Saf Health Care* 2004;13:223-5.
9. Lincoln YS, Lynham SA, Guba EG. Paradigmatic controversies, contradictions, and emerging confluences, revisited. *The Sage Handbook of Qualitative Research*. Vol. 4. Sage Publications; 2011. p. 97-128.
10. CASP. CASP Qualitative Checklist: Critical Appraisal Skills Program; 2013. Available from: http://www.media.wix.com/ugd/dded87_29c5b002d99342f788c6ac670e49f274.pdf. [Last cited on 2015 Mar 01].
11. Barbour RS. Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *BMJ* 2001;322:1115-7.
12. Popay 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.
13. Sale JE. How to assess rigour...or not in qualitative papers. *J Eval Clin Pract* 2008;14:912-3.
14. Meyrick J. What is good qualitative research? A first step towards a comprehensive approach to judging rigour/quality. *J Health Psychol* 2006;11:799-808.
15. Kitto SC, Chesters J, Grbich C. Quality in qualitative research. *Med J Aust* 2008;188:243-6.
16. Waterman AS. The humanistic psychology-positive psychology divide: Contrasts in philosophical foundations. *Am Psychol* 2013;68:124-33.
17. Finfgeld-Connett D. Generalizability and transferability of meta-synthesis research findings. *J Adv Nurs* 2010;66:246-54.
18. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Adm Policy Ment Health* 2013. p. 1-12.
19. Coyne IT. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? *J Adv Nurs* 1997;26:623-30.
20. Becker PH. Common pitfalls in published grounded theory research. *Qual Health Res* 1993;3:254-60.
21. Lincoln YS, Guba EG. *Naturalistic Inquiry*. Newbury Park, London: Sage Publications; 1985. p. 416.
22. Rodgers BL, Cowles KV. The qualitative research audit trail: A complex collection of documentation. *Res Nurs Health* 1993;16:219-26.
23. Kahn DL. Reducing bias. In: Cohen MZ, Kahn DL, Steeves RH, editors. *Hermeneutic Phenomenological Research: A Practical Guide for Nurse Researchers*. Newbury Park, London: Sage Publications; 2000. p. 85-99.
24. Carcary M. The research audit trail - Enhancing trustworthiness in qualitative inquiry. *Electron J Bus Res Methods* 2009;7:11-24. Available from: <http://www.ejbrm.com>. [Last accessed on 2015 Mar 03].
25. Jansen H. The logic of qualitative survey research and its position in the field of social research methods. *Forum Qual Soc Res* 2010;11:2.
26. Miles MB, Huberman AM. *Qualitative Data Analysis: An Expanded Sourcebook*. Newbury Park; London: Sage; 1994.
27. George M, Apter AJ. Gaining insight into patients' beliefs using qualitative research methodologies. *Curr Opin Allergy Clin Immunol* 2004;4:185-9.
28. Grossoehme DH. Overview of qualitative research. *J Health Care Chaplain* 2014;20:109-22.
29. Silverman D. *Doing Qualitative Research*. 3rd ed. Newbury Park, London: SAGE Publications Ltd.; 2009. p. 472.
30. Patton MQ. Enhancing the quality and credibility of qualitative analysis. *Health Serv Res* 1999;34:1189-208.
31. Allmark P. Popper and nursing theory. *Nurs Philos* 2003;4:4-16.
32. Kvale S, Brinkmann S. *Interviews: Learning the Craft of Qualitative Research Interviewing*. Newbury Park, London: Sage; 2009.
33. Trochim WM. *Research Methods: The Concise Knowledge Base*. Cincinnati, Ohio: Atomic Dog Publishing; 2005.
34. Zimmer L. Qualitative meta-synthesis: A question of dialoguing with texts. *J Adv Nurs* 2006;53:311-8.
35. Glaser BG, Strauss AL, Strutzel E. The discovery of grounded theory; strategies for qualitative research. *Nurs Res* 1968;17:364.
36. Van Manen M. *Researching Lived Experience: Human Science for An Action Sensitive Pedagogy*. New York. SUNY Press: Suny Press; 1990.
37. Noblit GW, Hare RD. *Meta-Ethnography: Synthesizing Qualitative Studies*. Newbury Park, London: Sage; 1988.
38. Gadamer HG. *Truth and Method* (Weinsheimer J, Marshall DG. Trans). New York: Continuum; 1989.
39. Thompson J. *Hermeneutic inquiry. Advancing Nursing Science Through Research*. Vol. 2. Sage Publications: Newbury Park, London; 1990. p. 223-86.
40. Toye F, Seers K, Allcock N, Briggs M, Carr E, Andrews J, *et al*. Trying to pin down jelly - Exploring intuitive processes in quality assessment for meta-ethnography. *BMC Med Res Methodol* 2013;13:46.
41. Rolfe G. Validity, trustworthiness and rigour: Quality and the idea of qualitative research. *J Adv Nurs* 2006;53:304-10.
42. Porter S. Validity, trustworthiness and rigour: Reasserting realism in qualitative research. *J Adv Nurs* 2007;60:79-86.
43. Guba EG, Lincoln YS. *Fourth Generationm Evaluation*. Newbury Park, London: Sage Publications; 1989. p. 296.
44. Sparkes AC. Myth 94: Qualitative health researchers will agree about validity. *Qual Health Res* 2001;11:538-52.

How to cite this article: Leung L. Validity, reliability, and generalizability in qualitative research. *J Family Med Prim Care* 2015;4:324-7.
Source of Support: Nil. **Conflict of Interest:** None declared.