

Appraising Qualitative Research for Evidence Syntheses: A Compendium of Quality Appraisal Tools

Qualitative Health Research
2018, Vol. 28(13) 2115–2131
© The Author(s) 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1049732318785358
journals.sagepub.com/home/qhr



Umair Majid¹  and Meredith Vanstone¹

Abstract

As the movement toward evidence-based health policy continues to emphasize the importance of including patient and public perspectives, syntheses of qualitative health research are becoming more common. In response to the focus on independent assessments of rigor in these knowledge products, over 100 appraisal tools for assessing the quality of qualitative research have been developed. The variety of appraisal tools exhibit diverse methods and purposes, reflecting the lack of consensus as to what constitutes appropriate quality criteria for qualitative research. It is a daunting task for those without deep familiarity of the field to choose the best appraisal tool for their purpose. This article provides a description of the structure, content, and objectives of existing appraisal tools for those wanting to evaluate primary qualitative research for a qualitative evidence synthesis. We then discuss common features of appraisal tools and examine their implications for evidence synthesis.

Keywords

quality appraisal; qualitative evidence synthesis; meta-synthesis; appraisal tool; systematic review

Background

A need for more holistic evidence to inform health policy decision-making has shone a spotlight on syntheses of qualitative evidence (Carroll, 2017; Lewin et al., 2018). The near-doubling of published qualitative evidence syntheses (QES) in the past decade (Hannes & Macaitis, 2012) reflects the push for clinical policies to include patient values, beliefs, and preferences (McInnes & Wimpenny, 2008). The rise of QES may be encouraged by the incentive for academic researchers to produce multiple publications and to borrow the evidentiary authority associated with claiming an article as a “synthesis” of multiple studies (Thorne, 2017). Reflecting the meta-analytic tradition, many QES authors choose to appraise the quality of studies included in their syntheses, with the intention of ensuring representation of the literature, assisting readers to evaluate the credibility of conclusions, and allowing decision-makers to understand the transferability of the findings (Whiting, Wolff, Mallett, Simera, & Savović, 2017). Quality appraisal is promoted as a step in the review process that allows researchers to reflect on the features of a research article that supposedly represent its methodological rigor and how the findings may inform health policy decision-making (Carroll & Booth, 2015). Some researchers believe that using a structured appraisal tool for quality assessment provides

an objective evaluation of the rigor of research (Whiting et al., 2017). Although contested by some qualitative researchers, appraisal tools used in multiple research paradigms advance the belief that the method and style of reporting key features of a study, including its theory, methodology, and methods, are critical to the quality appraisal process (Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009). In qualitative research, this may be accomplished through a rigid set of criteria in the form of an appraisal checklist (e.g., Lockwood, Munn, & Porritt, 2015), or through employing a more holistic approach by creating a reflexive dialogue between appraisers (e.g., Stige, Malterud, & Midtgarden, 2009). Many who choose to appraise the quality of qualitative research do so to ensure that the most methodologically sound studies are represented in the aggregation, integration, and synthesis of primary findings (Lewin et al., 2015). However, the relationship between the quality appraisal process and the methodological rigor of the resulting QES is complex.

¹McMaster University, Hamilton, Ontario, Canada

Corresponding Author:

Meredith Vanstone, Assistant Professor, Department of Family Medicine, David Braley Health Sciences Centre, McMaster University, 5003E, 100 Main Street West, Hamilton, ON, Canada L8P 1H6.
Email: Meredith.vanstone@mcmaster.ca

For example, an inadequate quality appraisal process may underrate or overrate the quality of an article, which may adversely affect the trustworthiness of findings from a study (Carroll, Booth, & Lloyd-Jones, 2012). Similarly, excluding a study on the basis of an appraisal verdict of low quality may prioritize certain types of studies, for instance, those with theoretically sophisticated findings. This may have the effect of de-prioritizing or limiting the impact of studies which may have important and useful descriptive findings, but which are not as theoretically sophisticated. Therefore, it is essential for QES researchers who engage in quality appraisal to ensure that the process they choose is rigorous and credible. This requires the selection of an appraisal tool that fits the aims and assumptions of the review.

Reflecting the diversity in qualitative traditions and approaches, there is no consensus on the necessity, merit, or appropriate approach to appraising the quality of qualitative research. To orient ourselves to the various positions one may take on quality appraisal, we have used Denzin's (2009) framework of three perspectives. Denzin's first perspective, the "foundationalists," argues that the same criteria for high-quality quantitative research should be applied to qualitative research. On the other side of the spectrum, "non-foundationalists" purport the inability of any criteria or prescriptive process to adequately judge the creativity and conceptual rigor of qualitative inquiry (Dixon-Woods, Shaw, Agarwal, & Smith, 2004). Downe (2008) describes this group as "believing that any externally imposed rules of acceptability for context-specific in-depth studies risks violating epistemological principles of knowledge as particular, specific, and resistant to exact replication" (p. 6). In between these two positions are the "quasi-foundationalists," who argue for the use of quality criteria that are unique to qualitative inquiry (Denzin, 2009). This perspective is most commonly adopted by QES researchers; it is the basis for the creation and use of appraisal tools for judging the rigor of qualitative research and efforts to standardize the methodological reporting of qualitative research. In designing and conducting this article, we draw upon the "quasi-foundationalist" perspective, believing that quality criteria developed for the purpose of appraising qualitative research can strengthen the understanding and applicability of QES. In our own work conducting QES for a variety of policy-making audiences, we have adopted a pragmatic approach, moving between non-foundationalism and quasi-foundationalism to reflect the purpose and function of the QES, as well as the traditions and preferences of the policy partner (Brundisini et al., 2015; Kandasamy, Khalid, Majid, & Vanstone, 2017; Majid, Kandasamy, Arora, & Vanstone, in press; Vanstone, Kandasamy, Giacomini, DeJean, & McDonald, 2016).

The number and diversity of available quality appraisal tools makes selection difficult even for experienced qualitative researchers. Navigating through over 100 appraisal tools is a cumbersome process, especially when encountering their diverse methodologies, philosophical perspectives, and purposes (Santiago-Delefosse, Gavin, Bruchez, Roux, & Stephen, 2016). Moreover, understanding the development of an appraisal tool, and its philosophy and purpose, becomes a strain on researchers engaging in a QES with limited time and resources, a common challenge for those working with policy partners. There are many reasons that authors of QES may choose to forego quality appraisal, including philosophical objections (Dixon-Woods et al., 2007) or time and resource limitations. Moreover, the difficulty of identifying a feasible tool that is congruent with the aims and approach of the synthesis may be an important reason why some QES reports may not use formalized quality criteria to appraise their studies (Dixon-Woods, Booth, & Sutton, 2007). Nevertheless, there is a growing trend toward the use of appraisal tools in QES (Hannes & Macaitis, 2012), which may depend on the purpose of research, type of QES, and its academic discipline or philosophical tradition (Carroll & Booth, 2015; Garside, 2014). Recently, GRADE-CERQual (Confidence in the Evidence From Reviews of Qualitative Research) published a series of papers with guidance on how to evaluate and use the findings of QES in the development of guidelines and policies (Lewin et al., 2018). These discussions show that the debate in many QES camps is shifting from whether the use of tools is appropriate to discussions of which tools are suitable for which purposes (Hannes & Macaitis, 2012).

The objective of this article is to identify and describe a variety of appraisal tools available for appraising primary qualitative research studies in the context of QES. By providing an overview of these tools, including their origin, purpose, content, structure, strengths, and criticisms, this article will assist researchers, decision-makers, and learners to choose an appraisal tool that best fits their purposes. After a description of the available tools, we offer an integrative, critical, and comparative analysis of these tools that addresses their historical antecedents, common patterns regarding structure and content, and the implications of these patterns on the QES process.

While the purpose of the current article is to assist researchers undertaking QES select a suitable appraisal tool, it is important to acknowledge the valid and important reasons why researchers may wish to abstain from formal quality appraisal. Due to the diversity of epistemological perspectives in qualitative research, there is a lack of consensus on the methods and standards for critical appraisal (Melia, 2010); some researchers are reluctant to evaluate the work of another researcher on a priori standards of quality or rigor. As we detail in the Discussion

section of this article, many qualitative researchers acknowledge that methodological and procedural details are typically underreported and their absence from a manuscript does not necessarily indicate their absence from the research process (Sandelowski & Barroso, 2007). Furthermore, in qualitative research the quality of findings often rests more on the conceptual prowess of researchers than it does on the creation and execution of a rigorous methodological process (Melia, 2010). Many types of QES authors are skeptical of independent assessments of quality and do not recommend the exclusion of studies on this basis (e.g., Barnett-Page & Thomas, 2009; Finfgeld-Connett, 2003; Noblit & Hare, 1988; Saini & Shlonsky, 2012; Sandelowski & Barroso, 2002, 2003). As a reflexive note, in our own QES work we have tried various forms of quality appraisal but have never excluded studies on the basis of quality appraisal. Consistent with Sandelowski and Barroso's (2007) qualitative meta-synthesis, we *have* excluded studies that offer no data or evidence to support their conclusions. We generally find that the absence of quality appraisal has not influenced our final conclusions and policy recommendations nor changed the way we represented our findings. This has been especially true in QES with large numbers of studies (Brundisini et al., 2015; Kandasamy et al., 2017; Majid et al., in press; Vanstone et al., 2016).

Method

By identifying and describing quality appraisal tools that are useful, relevant, and helpful for appraising individual qualitative research studies, we aim to offer those engaged with QES a resource to evaluate existing tools and decide which is the best fit consider their resources and QES objectives. To this end, rather than seek to describe an exhaustive list of all available appraisal tools for QES, we focused on identifying "high-utility" tools. We understand "high-utility" to describe tools that are some combination of available, familiar, authoritative, and easy-to-use tools that produce valuable results and offer guidance for their use. We chose to offer a more detailed description and comparison of these high-utility tools, rather than search for and describe all available tools, as many tools that exist simply combine other tools, may not have been used widely beyond the group that develops or supports them, or may not have been used specifically within a QES.

A clarifying note on terminology: In our search, we differentiated between appraisal tools and articles about appraisal tools. *Appraisal tools* were defined as instruments that enable QES researchers to perform the quality appraisal of a primary research study. Some appraisal tools may be classified as *checklists*, which provide a stepwise process with specific prompts to assess the quality of qualitative research. Other tools may be classified as *deliberative*

guides to evaluating the rigor of a qualitative study. *Articles about tools*, on the contrary, included papers that described the design, structure, and content of an appraisal tool or the validation of a tool in a QES. During our eligibility assessment, we observed that many of these articles modified the tool significantly (e.g., Campbell et al., 2003). To avoid confusion between the original and modified version of appraisal tools, we excluded articles that were validating a tool in a QES. For our analysis, we only included articles that described the design, structure, and content of appraisal tools. During our analysis, however, we did draw on these *articles about tools* to understand the strengths and criticisms of included tools.

We identified relevant, high-utility appraisal tools through a systematic literature search, building upon the work of Santiago-Delefosse and colleagues (2016) who performed a comprehensive and systematic search for qualitative appraisal tools in the literature. They identified 133 qualitative appraisal tools developed between 1985 and 2014 (specific date not available). They reduced their list of 133 tools to 58 because of redundancy in their structure, origin, and content. Santiago-Delefosse and colleagues (2016) may have excluded some of their 133 tools due to the difficulty in distinguishing between articles that describe the structure, design, and content of tools and articles that modified tools and validated them in a QES. In addition to the 58 tools, we extended their search to December 1, 2016, by replicating their search strategy in PsycINFO, MEDLINE, Embase, and HealthSTAR, which yielded 1,354 unique hits and four new tools not included in Santiago-Delefosse and colleague's (2016) list of 58 appraisal tools.

Sixty-two qualitative appraisal tools were assessed for eligibility according to criteria we designed to identify high-utility appraisal tools that had indicators of acceptability by the research community and adequate available information regarding their use and application in QES (Table 1). The purpose of our search and eligibility assessment was to identify and describe a few high-utility appraisal tools rather than to provide a comprehensive description of all the tools in the literature for qualitative appraisal. Eligible tools explicitly stated their suitability for the quality appraisal of qualitative studies in a QES. Tools were eligible if they declared this purpose in a supplementary resource or in the abstract, research objectives, discussion, or conclusions of their article. Appraisal for QES is not always the primary purpose of each tool, but if stated as a secondary or tertiary purpose, that tool was still eligible. For example, Booth and colleagues (2014) stated appraisal for QES as a secondary purpose of the Consolidated Criteria for Reporting Qualitative Studies (COREQ), a tool which we judged to be eligible as a high-utility tool due to its ease of use, broad acceptance, and wide adoption across disciplines. All included tools were accessible to the authors through the university

Table 1. Eligibility Criteria of High-Utility Tools.

Included	Excluded
<ul style="list-style-type: none"> • Tools published between 1985 and December 1, 2016. • Available in English • Explicit statement of usefulness for quality appraisal in QES • Employed in the appraisal process of a QES within the past 10 years (2006–2016) • Accessible to the authors via institutional access or request to corresponding author • Published in a peer-review journal or supported by an education, health care, or governmental institution 	<ul style="list-style-type: none"> • Tools not readily available in English • No explicit statement of relevance to QES • No citation evidence that this tool has been used for appraisal in the context of a QES within the past 10 years • Tools which require the collection of a fee for each use • Tools without academic or policy credibility as indicated by institutional support

Note. QES = qualitative evidence syntheses.

library system, the interlibrary loan system, or through correspondence with the listed authors. Eligible appraisal tools must have been published in a peer-reviewed journal or be supported by an organization affiliated with an educational, health care, or governmental institution (e.g., tools published on an institutional website). Finally, included appraisal tools must have been used in the quality appraisal process of a QES within the past 10 years, which was determined by searching who had cited that tool according to Google Scholar. With these criteria, we identified eight high-utility tools for appraising qualitative research for QES from the initial list of 62 tools. An appendix containing all 62 tools, their citations, and our eligibility assessment based on our predetermined criteria is provided as a supplementary document. Figure 1 shows our searching and sorting process.

Results

Eight appraisal tools were analyzed for their origin, purpose, content and structure, strengths, and criticisms. These features are organized in Table 2, which provides basic information about each tool including the author, year and country of origin, the primary discipline of the authors, and a description of the structure, content, and criteria of each tool. Table 2 also provides a summary of strengths and critiques of each tool. We have abbreviated the tools throughout the article as follows: CASP (Critical Appraisal Skills Programme, 2016), COREQ (Tong, Sainsbury, & Craig, 2007), ETQS (Long & Godfrey, 2004), JBI (Joanna Briggs Institute, 2016), Popay (Popay, Rogers, & Williams, 1998), QF (Spencer, Ritchie, Lewis, & Dillon, 2003), SRQR (O'Brien, Harris, Beckman, Reed, & Cook, 2014), and Walsh (Walsh & Downe, 2006).

Origin

The *origin* of the tool comprises the academic discipline, the citation of tool, and country of development. This information provides guidance for understanding the

context in which the tool was developed. Moreover, it may assist appraisers in referring to these tools for clarification on their purpose, content, and application in QES. Among the eight tools analyzed in this study, five (62.5%) identified with a biomedical discipline (COREQ, JBI, Popay, SRQR, Walsh), one (12.5%) was situated in sociology (ETQS), one (12.5%) in policy work (QF), and one (12.5%) did not identify with a particular discipline (CASP). Five (62.5%) were developed by authors primarily working in the United Kingdom (CASP, ETQS, Popay, QF, Walsh), whereas two (25%) were affiliated with Australia (COREQ, JBI) and one (12.5%) with the United States (SRQR).

Purpose

The *purposes* of assessing quality were identified iteratively upon review of the 62 appraisal tools assessed for eligibility. Appraisal of quality criteria can serve the following functions:

1. An augment to the quality appraisal of qualitative studies in QES
2. An educational tool for learners and novice qualitative researchers
3. A tool to streamline the appraisal of qualitative studies for health care professionals
4. A reporting standard for qualitative manuscripts for journal publication
5. A quality standard for specific types of research (e.g., rapid reviews, policy work, etc.)

Among the eight tools eligible for this analysis, all identified with the first function because this was an eligibility requirement for the tools included in this study. However, five tools had multiple purposes (CASP, COREQ, QF, SRQR, Walsh), and for two tools (COREQ, SRQR), No. 1 was not the primary objective. Most of the 62 tools assessed for eligibility explicitly identified with either purpose No. 3 (13%), No. 4 (47%), or both (7%).

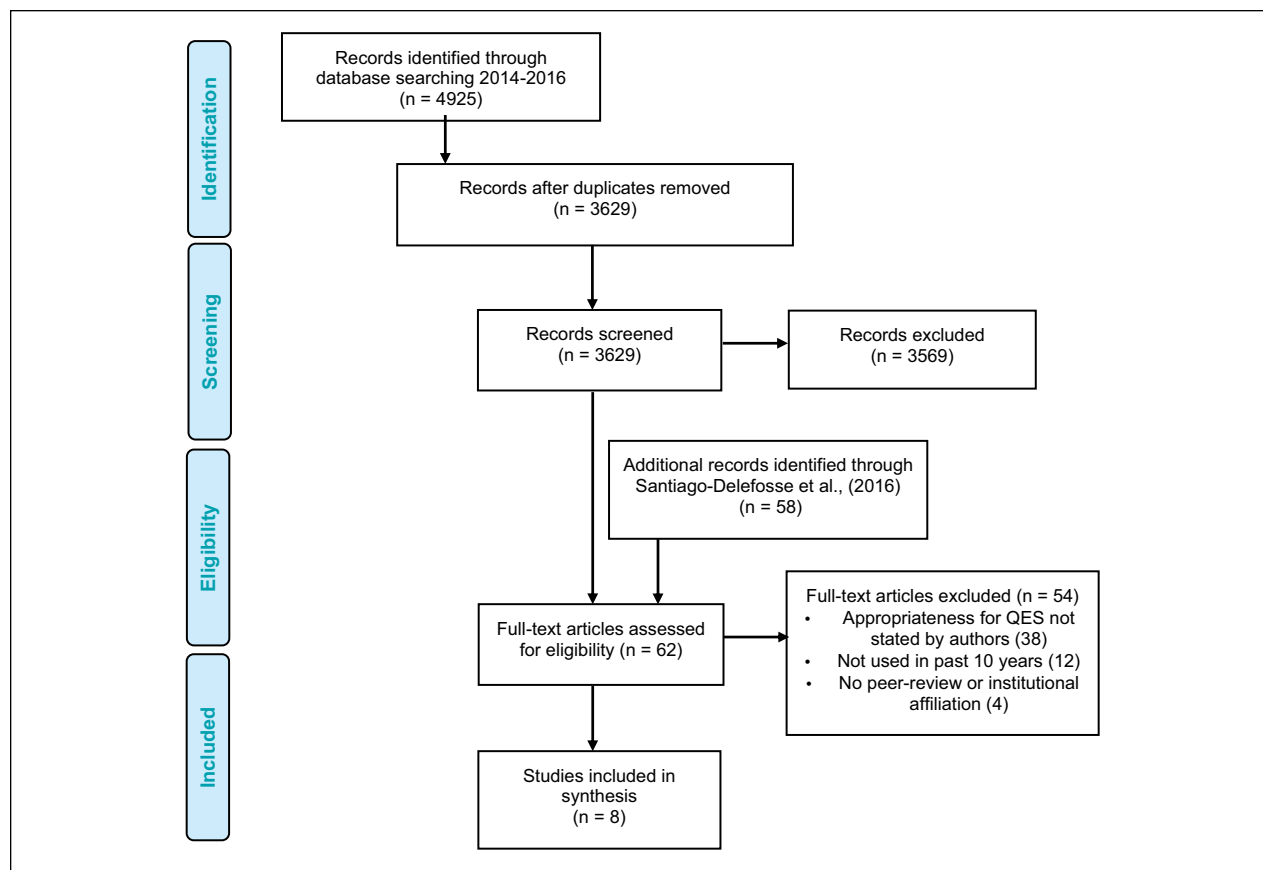


Figure 1. PRISMA flow diagram.

Source. Adapted from Moher, Liberati, Tetzlaff, Altman, and Prisma Group (2009).

Note. QES = qualitative evidence syntheses.

Content and Structure

The *content* and *structure* of appraisal tools is essential for understanding, applying, and comparing tools. The description provided in Table 2 stays close to the authors' explanations and organization of their tools. Criteria may be categorized based on whether they are in the form of questions (interrogative) or statements (declarative), and whether they elicit an open or closed response. For example, COREQ's Criterion #23 ("Transcripts returned—Were transcripts returned to participants for comment and/or correction") elicits a closed response.

Differentiating between interrogative and declarative statements is helpful in identifying appraisal tools that aim to create a reflexive dialogue between appraisers, and those that advance a firmer application of quality criteria. The majority of tools framed quality criteria as interrogative (CASP, COREQ, ETQS, JBI, Popay, QF), whereas some were framed as declarative statements (SRQR, Walsh).

Appraisal tools, on average, consisted of 19 quality criteria ranging from eight (Popay) to 86 (QF). Of the eight tools, five (62.5%) contained structured quality

criteria (CASP, COREQ, ETQS, QF, JBI), two of which (40%) were in the form of appraisal checklists (CASP, JBI). Information on whether the authors recommend the tool should be applied strictly or loosely is shown in Table 2. Some tools referred to additional documents containing either additional discussion on quality criteria (CASP; COREQ), guidelines and examples concerning the application of quality criteria to the appraisal process (SRQR), or additional information on the development of the tool (JBI).

Surprisingly, ethical concerns were not represented as a quality criterion in all the appraisal tools. In particular, two tools (COREQ; Popay) did not explicitly consider ethical issues as a quality criterion. Among the other six tools that considered ethics, three tools (CASP, JBI, SRQR) represented ethics as an appraisal item that inquired about the presence of statements of the study's approval from an institutional review board, informed consent, or the whether or not ethical issues were considered by the authors but not how or to what extent they were considered. The three remaining tools (ETQS, QF, Walsh) elicited a reflexive discussion about how the

Table 2. Descriptive Data of Qualitative Appraisal Tools.

Appraisal Tool	Country and Discipline	Purpose(s)	Content and Details	Strengths and Criticisms
CASP Qualitative Checklist, 2016 Critical Appraisal Skills Programme (CASP) 10 Questions to help you make sense of qualitative research: Milton Keynes Primary Care Trust, 2002.	United Kingdom N/A	<ol style="list-style-type: none"> 1. Appraisal in a qualitative evidence synthesis 2. Reporting standards for qualitative papers 3. Educational tool for learners 	<p>10 checklist questions answered with a yes, no or can't tell</p> <ul style="list-style-type: none"> • Screening Questions (2): Aims of study and appropriateness of qualitative methodology to aims • Appraisal Questions (8): Research design, recruitment strategy, data collection, reflexivity-related issues, ethical issues, rigor of data analysis, and the reporting and value of findings. <p>Press releases are available containing a discussion on quality criteria for qualitative research.</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Easy to understand and administer. • Easy to use as an educational tool for learners and novice researchers. • Most commonly used appraisal tool in the quality appraisal process of qualitative evidence syntheses. <p>Criticisms</p> <ul style="list-style-type: none"> • Weaker in evaluation of methodological quality compared to other appraisal tools (Hannes, Lockwood, & Pearson, 2010). • Adaptations of the tool are time-consuming to use (Campbell et al., 2003) • Favors studies with better methodological quality but may make weaker contributions to field (Dixon-Woods et al., 2007)
COREQ Tong, Sainsbury, and Craig (2007)	Australia, Medicine	<ol style="list-style-type: none"> 1. Reporting standards for qualitative papers 2. Appraisal in a qualitative evidence synthesis 	<p>32 structured statements in 3 domains:</p> <ul style="list-style-type: none"> • "Research Team and Reflexivity" (8): Researcher's personal characteristics and relationship with participants • "Study Design" (15): Theoretical framework, sampling strategy, study setting and data collection procedures • "Analysis of Findings" (9): Number of data coders, type of software used and method of reporting findings <p>Supplementary document published with Cochrane exploring issues with COREQ and quality for qualitative research.</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Widely used and endorsed by many international journals. • Developed through a systematic search and review of previously developed appraisal tools. <p>Criticisms</p> <ul style="list-style-type: none"> • Lack of emphasis on influence of philosophy on qualitative inquiry. • Only applicable to studies that employ focus groups and interviews as data collection methods. • Lack of guidelines to evaluate studies that use non-traditional data collection methods, or a combination of traditional and non-traditional methods.

(continued)

Table 2. (continued)

Appraisal Tool	Country and Discipline	Purpose(s)	Content and Details	Strengths and Criticisms
ETQS Long and Godfrey (2004)	United Kingdom, Sociology	I. Appraisal in a qualitative evidence synthesis	<p>38 unstructured questions under 4 themes:</p> <ul style="list-style-type: none"> • “Phenomenon Studied and Context” (17): Theoretical framework, setting, sample and outcomes • “Ethics” (3): Approval from a research ethics board, informed consent and reporting on ethical issues • “Data Collection, Analysis and Potential Research Bias” (10): Data collection process, analysis procedures and research bias • “Policy and Practice Implications” (4): Policy and practice implications of study findings <p>Evaluative Abstract (4): Summary of a study’s essential details and potential value (bibliographic details, key findings and summary of its strengths).</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Very comprehensive. • Evaluative abstract serves as a summary of the study. <p>Criticisms</p> <ul style="list-style-type: none"> • Lack of emphasis on philosophy and its congruity with methodology and methods. • Requires qualitative expert use. • More time-consuming than other appraisal tools.
JBI Lockwood, Munn, and Porritt (2015)	Australia, Medicine	I. Appraisal in a qualitative evidence synthesis	<p>10 checklist questions answered with a yes, no, unclear or not applicable:</p> <ul style="list-style-type: none"> • Congruity (5): Congruity between philosophical perspective, methodology, research question and objectives, data collection methods, data analysis and interpretation of results • Reflexivity (2): Research’s reflexivity and influence on the data collection and analysis procedures • Voice of Participants (1): Role of participants in the findings • Ethics (1): Approval from an ethics body and other ethical issues • Flow (1): Coherence between conclusions and interpretation of data <p>Supplementary document containing the methodology for development of JBI Critical Appraisal Checklist. Supplementary document describing the development and use of Qualitative Assessment and Review Instrument.</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Qualitative Assessment and Review Instrument (QARI) is an online software that streamlines the quality appraisal process using the JBI tool. • Brevity and clarity permits its use by less experienced researchers. • Better assessment of study details than other tools (Hannes, 2010). • Short and easy to use. <p>Criticisms</p> <ul style="list-style-type: none"> • Main emphasis is on congruity between philosophy, methodology and methods.

(continued)

Table 2. (continued)

Appraisal Tool	Country and Discipline	Purpose(s)	Content and Details	Strengths and Criticisms
Popay Popay, Rogers, and Williams (1998)	United Kingdom, Medicine	1. Appraisal in a qualitative evidence synthesis	<p>Eight “markers”:</p> <ul style="list-style-type: none"> Findings illuminate subjective meaning Iterative nature of research methods Appropriateness of sample Adequate description of findings Triangulation between data sources Flow of data, examples, interpretation Claims of generalizability <p>No information on scoring of markers.</p>	<p>Strengths</p> <ul style="list-style-type: none"> Seminal formulation of quality criteria for QES. Highlights the epistemological and ontological differences between qualitative and quantitative research. <p>Criticisms</p> <ul style="list-style-type: none"> Lack of information on the tool’s development methodology. Quality criteria have evolved and expanded since this work. No detailed guidelines on how to use and score this guide in the quality appraisal process making it difficult to report quality in a QES.
QF Spencer, Ritchie, Lewis, and Dillon (2003)	United Kingdom, Policy	<ol style="list-style-type: none"> Appraisal in a qualitative evidence synthesis Primary purpose is for use in evaluative and policy research 	<p>18 structured appraisal questions with 86 quality indicators categorized under 9 themes</p> <ul style="list-style-type: none"> Findings: Scope, explanation, contribution and credibility of findings Design: Defensibility of design Sample: Defensibility of sample Data Collection: Data collection process Analysis: Context and rigor of data analysis procedures Reporting: Link between data, interpretations and conclusions Reflexivity & Neutrality: Clarity of assumptions, perspectives and values Ethics: Evidence of attention to ethics Auditability: Adequate documentation of the research process <p>Each principle is associated with “quality indicators” that focuses on specific issues.</p>	<p>Strengths</p> <ul style="list-style-type: none"> Transparent guidelines and development process. Very comprehensive. <p>Criticisms</p> <ul style="list-style-type: none"> Very long and time-consuming. Lower reliability between appraisers than CASP (Dixon-Woods et al, 2007). Applies to only traditional data collection methods (i.e., interviews, documents, focus groups). Lack of guidelines on how to apply to distinct qualitative methodologies.

(continued)

Table 2. (continued)

Appraisal Tool	Country and Discipline	Purpose(s)	Content and Details	Strengths and Criticisms
O'Brien, Harris, Beckman, Reed, and Cook (2014)	United States, Academic Medicine	<ol style="list-style-type: none"> 1. Improve transparency of all aspects of qualitative research 2. Provide clear standards for reporting qualitative research 3. Assist the review process of evaluating a manuscript for potential publication 4. Appraisal in a qualitative evidence synthesis 	<p>21 structured statements into five themes:</p> <ul style="list-style-type: none"> • "Title and Abstract" (2): Concise description and summary • "Methods" (11): Qualitative approach, reflexivity, context, sampling, ethical issues, data collection methods and instruments, units of study, data processing and analysis and techniques to enhance trustworthiness • "Results/Findings" (2): Synthesis, interpretation and connections to data • "Discussion" (2): Incorporation of prior work, implications on the field, transferability and limitations of study <p>Supplementary document containing examples of applying criteria is available.</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Very transparent in purpose and development of appraisal tool. • Developed through a systematic search and review of previously developed appraisal tools. • Claims to consider all qualitative research methodologies equally. <p>Criticisms</p> <ul style="list-style-type: none"> • Nascent appraisal tool. • Whether it truly considers all qualitative methodologies equally, as suggested, is unestablished.
Walsh and Downe (2006)	United Kingdom, Medicine	<ol style="list-style-type: none"> 1. Appraisal in a qualitative evidence synthesis 2. Primary purpose is for use in rapid reviews in policy 	<p>12 "essential criteria" statements under eight stages of research process:</p> <ul style="list-style-type: none"> • Scope and Purpose: Study rationale, research question and existing literature • Design: Study design and data collection strategies. • Sampling Strategy: Eligibility criteria and appropriateness for methodology • Analysis: Coherence of analytical approach with sampling and design • Interpretation: Adequate description of context, audit trail and coherence between data and conclusions • Reflexivity: Sufficient description of how researcher may have influenced the findings • Ethical Dimensions: Researchers demonstrated concern about risk to human participants • Relevance and Transferability: Extent to which findings can be applied to different contexts <p>These criteria are to be used "imaginatively" instead of prescriptively, and therefore, rating of criteria is not recommended.</p>	<p>Strengths</p> <ul style="list-style-type: none"> • Use of "berry-picking" approach to find existing appraisal tools provides a real-world perspective where finding one tool online leads to finding another tool. This approach mimics how appraisers would locate tools when appraising the quality of a qualitative paper. <p>Criticisms</p> <ul style="list-style-type: none"> • Berry-picking approach is non-systematic and therefore not representative of the literature. • Limited information on details on development methodology. • Use of criteria "imaginatively" instead of prescriptively makes it difficult to report quality in a QES.

Note. QES = qualitative evidence syntheses.

study's procedures may harm or benefit the participants. The absence of ethical issues as a quality criterion may not mean that ethical issues are not considered in the appraisal tool. Rather, some tools may utilize the methodological quality of a research study as a proxy for whether or not a research study is utilizing its resources ethically and efficiently to investigate the phenomena. However, we believe that a few items that examine the ethical conduct of the research study are necessary as criteria in appraisal tools. QF, for example, provides a comprehensive assessment of ethical issues with seven quality indicators. Moreover, this tool uses words such as "evidence," "documentation," and "discussion" to elicit an open-ended conversation among appraisers about the ethical conduct of research. The way QF appraises ethical issues of a research study may serve as an example for other appraisal tools.

Strengths and Criticisms

The *strengths* and *criticisms* column provides appraisers with practical information concerning the application of appraisal tools in QES. These critiques derive from the literature on the application of quality criteria in qualitative research (e.g., Hannes & Macaitis, 2012), published manuscripts on the comparison and evaluation of tools included in this study (e.g., Masood, Thaliath, Bower, & Newton, 2011), identified patterns of quality appraisal in QES from the literature (e.g., Hannes, Lockwood, & Pearson, 2010), and our own expertise with quality appraisal and qualitative research. The objective of this assessment was to gather the insight of researchers experienced in the quality appraisal of qualitative research to represent the strengths and weaknesses of each appraisal tool. Some tools had more strengths or criticisms than others, which may be due to the quantity and quality of information available in the literature. For example, the CASP Qualitative Checklist is a commonly used appraisal tool and therefore more insight is available on its strengths, shortcomings, and limitations (Dixon-Woods, Sutton, et al., 2007).

A frequently reported criticism pertains to the length of appraisal tools. Many tools included in this study were brief, which is advantageous for application in health care and policy evaluations that require the efficient translation of evidence. Other appraisal tools were comprehensive, and their use requires a significant investment of time from multiple researchers with extensive expertise in qualitative methodology. Among the tools included in this study, we categorized three of eight (37.5%) as short (CASP, JBI, Popay), three of eight (37.5%) as moderate length (COREQ, SRQR, Walsh), and two of eight (25%) as long (ETQS, QF). This judgment was based on not only the number of items but also the number of themes,

subthemes, accompanying considerations, and additional factors relevant to using the tool. This discordance in the development and application of quality criteria is an important consideration for situating the appraisal process within the QES enterprise. We also note concern about the limited information on the development of appraisal tools (Popay, Walsh), and incomplete guidelines for their application in quality assessment (COREQ, Popay, QF). In terms of strengths, two notable considerations were tools developed through a systematic consideration of existing tools (COREQ, SRQR) and tools that emphasized transparency in the tool development and reporting processes (QF, SRQR).

Appraisal Tool Guide

The *Appraisal Tool Guide* (Table 3) is a comparison of appraisal tools. It is intended as a resource for researchers, decision-makers, and learners looking to appraise the qualitative studies in their QES to reflect on some features of appraisal tools that best fit their scholarly purpose. To assist tool-searchers in this task, Table 3 organizes tools according to frequently asked questions and frequently sought-after features. These judgments consider the structure, explicit purposes, identified patterns in QES from the literature, and our own analysis of the tools included in this study.

The Appraisal Tool Guide (Table 3) is not meant to serve as an exhaustive categorization of all potentially relevant features of quality appraisal tools. Rather, this guide intends to demarcate and explicate some features of tools that were most salient from our analysis. These features are represented in this guide in a way that assists researchers to consider the various strengths of each tool and how these strengths align with the purposes of their investigation. Of course, choosing a tool may require some compromises in priorities. The comprehensive tools are not those which facilitate rapid appraisal within the context of time and resource constraints. This is a natural dilemma of tool selection. We have designed Table 3 to identify several primary and secondary examples of each criterion, so that investigators facing a dilemma may consider additional values and ideally find a tool which satisfies multiple priorities. These dilemmas, although inconvenient, may force QES investigators to explicate their purpose, values, and objectives of both their scholarly investigation and the quality appraisal process. While the ultimate objective may be a guide that streamlines the process of deciding the most optimal tool for a particular study, we are not sure that there will ever be a single correct tool for any project; investigators may need to look for the "best fit." We hope this guide will create space for investigators to consider the various strengths of each tool against an explicit consideration of

Table 3. An Appraisal Tool Guide.

Question	Recommended Tool	Other Options
Who am I?		
Are you a novice researcher or learner?	CASP	
Are you a decision-maker?	CASP	JBI QF
Are you an experienced qualitative researcher?	Personal Judgment ^a or SRQR	
Length		
Are you looking for a tool that is short and easy-to-follow?	CASP	JBI Popay
Are you looking for a tool that is comprehensive?	QF	ETQS
Are you looking for a tool that balances brevity and comprehensiveness?	SRQR or COREQ	Walsh
Development		
Are you looking for a tool developed through a systematic literature review?	COREQ	SRQR
Are you looking for a tool developed through expert debate and discussion?	ETQS	Popay
Are you looking for a tool that used both literature review and expert debate with researchers?	SRQR	Walsh JBI QF
Application		
Are you looking for a tool that is specific to studies that have employed focus groups and interviews?	COREQ	
Are you looking for a tool that applies to all qualitative methodologies?	SRQR	JBI
Are you looking for a tool that is specifically designed for policy evaluations?	QF	
Are you looking for a tool that is most commonly used in QES?	CASP	JBI
Philosophical perspective		
Are you looking for a tool that is positivist or post-positivist?	CASP	COREQ ETQS
Are you looking for a tool that is constructivist?	SRQR	Walsh Popay
Other		
Are you looking for a tool that emphasizes the philosophy of the study and researcher, and its congruity with methodology and methods?	JBI	Popay
Are you looking for a tool that emphasizes reflexivity by the researcher?	JBI	Popay
Are you looking for a tool that highlights the epistemological and ontological differences between qualitative and quantitative research?	Popay	
Are you looking for a tool that focuses on the “integrity, transferability and transparency” of findings?	Walsh	QF
Do you require a summary of each study you appraise?	ETQS	

Note. QES = qualitative evidence syntheses.

^aPersonal judgment refers to the use of researcher’s own qualitative expertise to evaluate the rigor of a study. Some authors have found personal judgment to be superior to the use appraisal tools for the assessment of qualitative research (Dixon-Woods et al., 2007).

the aims and priorities of their QES project. Through this process, there is opportunity for researchers to deeply reflect on the nuances of quality appraisal as it pertains to their research investigation.

Discussion

This section provides an integrative and comparative discussion about the QES quality appraisal process deriving from the literature and our observations of the structure, content, purposes, and features of the appraisal tools included in this study. We have chosen to provide

an integrative discussion of three central controversies in qualitative appraisal scholarship: Methodological Reporting, Philosophical Foundations, and Reflexivity. We identified these three topics as central to both the choice to appraise, the selection of an appraisal tool and the employment of that tool. With this integrative discussion, we aim to provide readers some grounding in the guiding principles and issues that underpin scholarship in this area, so they may identify their own standpoint on these issues, formulate their own objectives for appraisal within QES and select a tool which best fits these values.

We begin this section with a discussion about how the style and content of the methodological details reported in a research study influences the quality assessment process when an appraisal tool is used. As a part of this discussion, we describe different perspectives to the use of tools in the quality assessment of qualitative research. On one hand, some scholars recommend the exclusion of articles based on a low-quality assessment, other scholars, on the other hand, contest this view because of the belief that quality appraisal tools may not truly assess the meaningfulness, utility, or potential impact of qualitative findings (Melia, 2010).

Still under debate is the extent to which appraisal tools can assess the influence of a study's philosophical orientation on the study findings. Among the appraisal tools included in our analysis, we observed a spectrum of tools with varying levels of emphasis on the role of philosophy in the quality assessment of qualitative research. In this section, we explore how the over- or under-emphasis in appraisal tools influences how QES quality appraisal is carried out. This discussion leads into the final topic where we examine the use reflexivity as a de facto quality criterion in all the appraisal tools analyzed in this study. Since appraisal tools may be challenged to evaluate the extent to which reflexivity used in primary research is useful, relevant, and appropriate, this section will elaborate on the challenges associated with using and assessing reflexivity with appraisal checklists followed by a brief overview of the implications for health policy and QES of using reflexivity as a universal quality criterion.

Methodological Reporting

There is an implicit notion among some qualitative scholars that inadequately reported study details are an indicator of low-quality findings that will undermine the integrity of the findings of the QES, and therefore, must be excluded from a QES (Carroll & Booth, 2015). However, there is no consensus on this notion. Carroll and colleagues (2012), for example, compared the interpretation of a QES with and without studies that were reported inadequately as determined by an appraisal checklist. They found minimal changes in the thickness, material, and depth of the QES, a finding shared by other studies (Noyes & Popay, 2007; Thomas & Harden, 2008). Others have encouraged the exclusion of a "weak study" that "stand[s] out as an anomaly" when compared with other studies in a QES (Estabrooks, Field, & Morse, 1994, p. 508). These scholars advocate for the exclusion of methodologically poor studies because they do not significantly or meaningfully contribute to the QES interpretation. On the contrary, some scholars have contested this view by asserting that inadequately reported studies may still provide meaningful contributions (Garside, 2014; Sandelowski, 2000) and if a qualitative

study "offers some theoretical insights, it is doing its job" (Melia, 2010, p. 572). Some of these scholars suggest that the exclusion of studies based on a priori criteria may adversely influence the generalizability of the QES (Carroll & Booth, 2015). Other scholars advocate for the inclusion of "soundly based findings . . . even when other findings from the same article might be rejected" (Jensen & Allen, 1996, p. 558). These scholars emphasize that inadequacy in reporting methodological details or findings in a manuscript does not mean that there is inadequacy in the conduct of research. Moreover, these scholars believe that the way the authors present the methodological details and findings of a study is not standard and may vary according to, among other factors, their academic discipline (Sandelowski & Barroso, 2006), experience in the qualitative research paradigm (Carroll & Booth, 2015; Thorne, 2017), analytic prowess (Giacomini & Cook, 2000), and the year of publication since recently published articles tend to have richer interpretations (Carroll et al., 2012).

The use of an appraisal checklist to exclude articles in a QES has raised concerns about whether it truly assesses the design and execution of the study or simply the quality of methodological reporting (Goldsmith, Bankhead, & Austoker, 2007). These concerns become more acute when an appraisal tool is used to exclude studies from a synthesis, rather than provide a context to the reader about the methodological quality of studies included in the QES. The CASP Qualitative Checklist, for example, due to its structured nature and checklist format, may focus its evaluation of quality on how qualitative research studies are reported in their journal articles rather than on evidence of analytic rigor, originality, or scholarly contribution to the field. Related to structured quality criteria is how the focus/purpose of an appraisal tool (e.g., for QES, for journal manuscripts, etc.) influences its design and content. Tools that are intended for streamlining the appraisal process for health care providers may emphasize structured quality criteria in an effort to accommodate time constraints. As a consequence of a more structured appraisal tool, the quality appraisal process may focus more on the reporting of methodological procedure rather than the value, contribution, and impact of the work. This can be troubling. If an appraisal tool is used to exclude articles based on prescriptive criteria, Thorne (2017) asserts that "the reader has no capacity to judge what gorgeous but imperfect interpretations may have been excluded, and what technically correct but 'bloodless' and unimaginative findings may have been privileged in delineating the final meta-synthesis sample" (p. 7).

Using checklist items to determine the rigor of qualitative research may increase the risk of the "tail wagging the dog" (Barbour, 2001), and increasing the use of "shorthand descriptions of credibility" such as triangulation to convey a false representation of rigor (Barbour &

Barbour, 2003, p. 181). Ethics is an example of a “checklist” item that may require deeper reflection and dialogue than the simple appraisal of presence or absence of a statement noting that institutional ethical approval was received. Including ethics as a checklist item may not obligate the appraiser to consider the indications of whether or not the research was conducted in a way that harms participants or other groups. This is significant, as institutional research ethics approval does not always indicate the ethical conduct of research. For example, research studies about Indigenous peoples may be reviewed and approved by an institutional ethics board and ergo fulfill the ethical quality criterion but these studies may still not exemplify a level of sensitivity to the particular ethical issues that surround Indigenous research, such as the need to consider the relationship of Indigenous people to their cultural knowledge, data, and information (First Nations Governance Center, 2014).

While checklists have many disadvantages, they also may equip novice qualitative researchers with the resources to evaluate qualitative research efficiently. Appraisal checklists may also serve as excellent springboards for learning and communicating the methodological practices and nuances of qualitative research to scholars from other research traditions (Belgrave, Zablotsky, & Guadagno, 2002). However, qualitative research is not “a descriptive science but also relies on the capacity to evoke imaginative experience and reveal new meanings” (Yardley, 2008, p. 260), making the creation and application of prescriptive and rigid quality criteria contradictory and practically discordant (Garside, 2014). Checklists should not be judiciously applied as authoritative determinants of quality in QES because they de-emphasize the larger interpretive context of research studies, which is at the heart of what qualitative research contributes to evidence-based practice and policy. For example, the JBI tool does not explicitly inquire about the value and relevance of each qualitative study in relation to the aim and direction of the QES. Not inquiring about value and relevance may result in the de-prioritization of the interpretive context of each study. The use of checklists in general may de-emphasize the larger interpretive context because their structured format may force appraisers to focus on specific procedural aspects of a qualitative study instead of encouraging appraisers to acknowledge, clarify, and reflect upon the various contexts that may influence the research study.

The Emphasis of Philosophy in Appraisal Tools

The discernment of a study’s philosophical orientation may be a challenging judgment to make because many qualitative studies do not explicitly report their philosophical orientation (McInnes & Wimpenny, 2008). For appraisal tools that rely heavily on these characteristics,

for example, the JBI Critical Appraisal Checklist, studies that do not explicitly state these characteristics may be assessed as low quality, which may be reflective of their reporting choices rather than the conduct of research. In some cases, appraisers may be able to ascertain implicit philosophical assumptions and the congruency of methodological decisions based on the information presented in the article. For example, the first five questions of JBI inquire about the congruency between the philosophy, methodology, and methods of a research study. This approach to appraisal, however, is more time-consuming and requires a reviewer who has knowledge and experience with the philosophical foundations of qualitative research. Where some tools rely heavily on the philosophical orientation of a study for quality assessment, other tools, such as the CASP Qualitative Checklist, place less importance on these characteristics. An alternative approach is exemplified in tools which inquire about a theoretical framework of a research study. The ETQS, for example, asks three questions to appraisers relevant to the theoretical framework: (a) What theoretical framework guides or informs the study? (b) In what ways is the framework reflected in the way the study was done? (c) How do authors locate the study within the existing knowledge base?

While it may be true that overlooking a study’s philosophical orientation may result in a less meaningful synthesis, overemphasizing these characteristics in the quality appraisal process may inadvertently value theoretically sophisticated qualitative articles over applied health research. This is troubling because when working to inform health policy, theoretical sophistication may not be necessary for a meaningful research contribution (Sandelowski, 2000). Descriptive research that stays close to qualitative data can be a useful approach to representing patient perspectives in qualitative research. This type of research may focus on qualitative findings that are the most salient, common, or conceptually powerful and which may be very relevant to health policy and practice. Therefore, the focus on theoretically dense qualitative articles over others may be problematic for a QES that aims to synthesize the best available evidence but misses or de-emphasizes applied, descriptive research that could make a valuable contribution to the policy question at hand.

Reflexivity as a Universal Quality Criterion?

Reflexivity—which may be understood as a methodological technique that explicates how the researcher may have influenced the research process—has become a well-established component of rigor in qualitative research (Koch & Harrington, 1998). Today, many decision-makers, grant agencies, and journal editors

use reflexivity as a de facto quality criterion to judge the quality of a qualitative research study. The appreciation of reflexivity as a way to establish the credibility of qualitative research reflects the constructivist or interpretivist methodological stances which emphasize the role of the investigator as a research instrument.

Reflexivity is commonly operationalized in the majority of qualitative appraisal tools. This process forces prescription onto a form of inquiry that is antithetical to standardization and instead driven by creativity, insight, and prowess. Reflexivity is increasingly evaluated using checklist items that require an appraiser to make value judgments from their own preconceptions of qualitative research. All tools included in this analysis acknowledged, albeit in different ways, the role of the researcher and participants in the research process, indicating a consensus concerning the use of reflexivity as a quality criterion. However, appraisal tools are often challenged to address the complexity of employing useful, relevant, and appropriate reflexive practices.

Assessing reflexivity is difficult because it may be interpreted in a multitude of ways depending on the discipline and jurisdiction of the appraiser (Gentles, Jack, Nicholas, & McKibbin, 2014). This concern is further intensified because historically, reflexivity has been loosely reported in qualitative articles with “most researchers failing to specify their understandings, positions, and approaches, ignoring how widely reflexivity has been conceptualized and the divergent ways it can be practiced” (Gentles et al., 2014, p. 2). The inappropriate application of reflexive thinking has notable implications for its use as a quality criterion because it is “akin to entering uncertain terrain where solid ground can all too easily give away to swamp and mire” (Finlay, 2002, p. 212). Moreover, the inadequate reporting of reflexivity makes it difficult to develop and apply quality criteria that assess the reflexive practices of a research study. How can a reviewer comprehensively assess whether the researcher of a primary study has engaged in reflexive thinking? What are the markers that indicate whether the reflexive thinking in a particular study is useful to the research study? What level of experience and expertise is required to make this value judgment? The appraisal tools included in this study require the quality assessment of reflexivity but do not provide guidance on how appraisers might handle these pertinent issues. To that end, appraisal tools may not be able to answer such questions about the usefulness of reflexivity in an investigation. Instead, reflexivity may be best judged through an open, collaborative discussion among researchers.

The quality appraisal process also requires appraisers to judge the extent to which the researcher of a primary qualitative study engaged in meaningful, reflexive thinking, and how the researcher may have influenced the

research process. However, this approach to rigor assessment using reflexivity is not straightforward because it is a subjective experience, and more broadly, a reflexive phenomenon in and of itself. The judgment is contingent on the appraiser’s philosophy and their beliefs about reality, truth, and knowledge. Other considerations that influence the assessment of reflexivity are the style and depth of reflexivity that may depend on “pragmatic considerations, such as the intended audience for published research” (Finlay, 2002, p. 227) or the limitations of the publication venue.

Because reflexivity is a component of most quality guidelines for qualitative research, studies that do not operationalize it in a way that is congruent to an appraisal tool may receive a negative quality assessment because of a mismatch between the requirements of the tool and the boundaries of what can be reported about the reflexivity of a research study. The ETQS, for example, frames reflexivity as “Researcher’s potential bias.” The use of “bias” in this tool may have derived from positivist traditions that may imply that the researcher’s assumptions adversely influence the investigation. The way reflexivity is framed in ETQS may be incongruent with a research study that views the assumptions of investigators as something that does not adversely influence the results but enhances and supplements the final interpretation and synthesis. Moreover, due to the ways in which reflexivity is operationalized in some appraisal tools, some efforts by researchers to explicate their reflexive thinking may not be appraised as being relevant, useful, or valuable to the research study. “Researchers are, in effect, damned if they do and damned if they don’t” (Finlay, 2002, p. 227).

Strengths and Limitations of This Study

This research provides an overview of key perspectives and stances in the field of quality appraisal of qualitative research, designed to inform researchers engaging in QES who may wish for an orientation to the scholarship in this area. Using a systematic search process, we identified eight high-utility tools and described these in sufficient depth so as to permit a QES team to identify and describe a tool that is the best fit for their purpose, resources, and philosophical commitments. In our focus on identifying high-utility tools which would be most useful to researchers searching for an approach to quality appraisal, we may have neglected to include other valuable tools which have not yet been broadly taken up, or used in the QES context. As the field of QES continues to evolve, no doubt additional tools will be developed, refined, and be recognized for their valuable contributions to quality appraisal.

Conclusion

This article reviewed the characteristics of appraisal tools designed for use in QES. We described eight appraisal tools that QES researchers may wish to consider, comparing their origin, purpose, content, strengths, and criticisms. The development of a consensus among qualitative researchers on appropriate quality appraisal tools is challenged by the wide variety of perspectives on reporting conventions, philosophical congruence, and reflexivity. The decision of which appraisal tool to use depends on the objectives of the evidence synthesis, the expertise of the researchers, and the time and resources available.

Acknowledgments

We appreciate the support and critical feedback from Dr. Michael Wilson of the McMaster Health Forum at McMaster University and the Department of Health Evidence and Impact at McMaster University.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.


Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: M. Vanstone receives salary support from the Government of Ontario and the Ontario SPOR SUPPORT Unit, which is supported by the Canadian Institutes of Health Research and the Government of Ontario.

Supplementary Material

Supplemental Material for this article is available online at journals.sagepub.com/home/qhr. Please enter the article's DOI, located at the top right hand corner of this article in the search bar, and click on the file folder icon to view.

ORCID iD

Umair Majid  <https://orcid.org/0000-0002-4581-7714>

References

- Barbour, R. S. (2001). Checklists for improving rigour in qualitative research: A case of the tail wagging the dog? *British Medical Journal*, *322*(7294), Article 1115.
- Barbour, R. S., & Barbour, M. (2003). Evaluating and synthesizing qualitative research: The need to develop a distinctive approach. *Journal of Evaluation in Clinical Practice*, *9*, 179–186. doi:10.1046/j.1365-2753.2003.00371.x
- Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: a critical review. *BMC medical research methodology*, *9*(1), 59. <https://doi.org/10.1186/1471-2288-9-59>
- Belgrave, L. L., Zablotsky, D., & Guadagno, M. A. (2002). How do we talk to each other? Writing qualitative research for quantitative readers. *Qualitative Health Research*, *12*, 1427–1439. doi:10.1177/1049732302238753
- Booth, A., Hannes, K., Harden, A., Noyes, J., Harris, J., & Tong, A. (2014). COREQ (Consolidated Criteria for Reporting Qualitative Studies). In D. Moher, D. G. Altman, K. F. Schulz, I. Simera, & E. Wager (Eds.), *Guidelines for reporting health research: A user's manual* (pp. 214–226). Oxford, UK: John Wiley.
- Brundisini, F., Vanstone, M., Hulan, D., DeJean, D., & Giacomini, M. (2015). Type 2 diabetes patients' and providers' differing perspectives on medication nonadherence: A qualitative meta-synthesis. *BMC Health Services Research*, *15*(1), Article 516. doi:10.1186/s12913-015-1174-8
- Campbell, R., Pound, P., Pope, C., Britten, N., Pill, R., Morgan, M., & Donovan, J. (2003). Evaluating meta-ethnography: A synthesis of qualitative research on lay experiences of diabetes and diabetes care. *Social Science & Medicine*, *56*, 671–684. doi:10.1016/S0277-9536(02)00064-3
- Carroll, C. (2017). Qualitative evidence synthesis to improve implementation of clinical guidelines. *British Medical Journal*, *356*, Article j80. doi:10.1136/bmj.j80
- Carroll, C., & Booth, A. (2015). Quality assessment of qualitative evidence for systematic review and synthesis: Is it meaningful, and if so, how should it be performed? *Research Synthesis Methods*, *6*, 149–154. doi:10.1002/jrsm.1128
- Carroll, C., Booth, A., & Lloyd-Jones, M. (2012). Should we exclude inadequately reported studies from qualitative systematic reviews? An evaluation of sensitivity analyses in two case study reviews. *Qualitative Health Research*, *22*, 1425–1434. doi:10.1177/1049732312452937
- Critical Appraisal Skills Programme. (2016). *CASP qualitative research checklist* [online]. Retrieved from <http://www.casp-uk.net/casp-tools-checklists>
- Denzin, N. K. (2009). The elephant in the living room: Or extending the conversation about the politics of evidence. *Qualitative Research*, *9*, 139–160. doi:10.1177/1468794108098034
- Dixon-Woods, M., Booth, A., & Sutton, A. J. (2007). Synthesizing qualitative research: A review of published reports. *Qualitative Research*, *7*, 375–422. doi:10.1177/1468794107078517
- Dixon-Woods, M., Shaw, R. L., Agarwal, S., & Smith, J. A. (2004). The problem of appraising qualitative research. *Quality and Safety in Health Care*, *13*, 223–225. doi:10.1136/qshc.2003.008714
- Dixon-Woods, M., Sutton, A., Shaw, R., Miller, T., Smith, J., Young, B., . . . Jones, D. (2007). Appraising qualitative research for inclusion in systematic reviews: A quantitative and qualitative comparison of three methods. *Journal of Health Services Research & Policy*, *12*, 42–47. doi:10.1258/135581907779497486
- Downe, S. (2008). Metasynthesis: A guide to knitting smoke. *Evidence Based Midwifery*, *6*(1), 4–9.
- Estabrooks, C. A., Field, P. A., & Morse, J. M. (1994). Aggregating qualitative findings: An approach to theory development. *Qualitative Health Research*, *4*, 503–511. doi:10.1177/104973239400400410
- Finlay, L. (2002). Negotiating the swamp: The opportunity and challenge of reflexivity in research practice. *Qualitative Research*, *2*, 209–230. doi:10.1177/146879410200200205

- Finfeld, D. L. (2003). Metasynthesis: The state of the art—so far. *Qualitative health research*, 13(7), 893-904. <https://doi.org/10.1177/1049732303253462>
- First Nations Governance Centre. (2014). *Ownership, Control, Access and Possession (OCAP TM): The path to first nations information governance*. Ottawa: First Nations Information Governance Centre. Retrieved from http://fnigc.ca/sites/default/files/docs/ocap_path_to_fn_information_governance_en_final.pdf
- Garside, R. (2014). Should we appraise the quality of qualitative research reports for systematic reviews, and if so, how? *Innovation: The European Journal of Social Science Research*, 27, 67–79. doi:10.1080/13511610.2013.777270
- Gentles, S. J., Jack, S. M., Nicholas, D. B., & McKibbin, K. (2014). Critical approach to reflexivity in grounded theory. *The Qualitative Report*, 19(44), 1–14. Retrieved from <https://nsuworks.nova.edu/tqr/vol19/iss44/3>
- 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. doi:10.1001/jama.284.3.357
- Goldsmith, M. R., Bankhead, C. R., & Austoker, J. (2007). Synthesising quantitative and qualitative research in evidence-based patient information. *Journal of Epidemiology and Community Health*, 61, 262–270. doi:10.1136/jech.2006.046110
- Hannes, K., Lockwood, C., & Pearson, A. (2010). A comparative analysis of three online appraisal instruments' ability to assess validity in qualitative research. *Qualitative Health Research*, 20, 1736–1743. doi:10.1177/1049732310378656
- Hannes, K., & Macaitis, K. (2012). A move to more systematic and transparent approaches in qualitative evidence synthesis: Update on a review of published papers. *Qualitative Research*, 12, 402–442. doi:10.1177/1468794111432992
- Jensen, L. A., & Allen, M. N. (1996). Meta-synthesis of qualitative findings. *Qualitative Health Research*, 6, 553–560. doi:10.1177/104973239600600407
- Joanna Briggs Institute. (2016). *Checklist for qualitative research* [online]. Retrieved from <http://joannabriggs.org/research/critical-appraisal-tools.html>
- Kandasamy, S., Khalid, A. F., Majid, U., & Vanstone, M. (2017). Prostate cancer patient perspectives on the use of information in treatment decision-making: A systematic review and qualitative meta-synthesis. *Ontario Health Technology Assessment Series*, 17, 1–32.
- Koch, T., & Harrington, A. (1998). Reconceptualizing rigour: The case for reflexivity. *Journal of Advanced Nursing*, 28, 882–890. doi:10.1046/j.1365-2648.1998.00725.x
- Lewin, S., Booth, A., Glenton, C., Munthe-Kaas, H., Rashidian, A., Wainwright, M., . . . Noyes, J. (2018). Applying GRADE-CERQual to qualitative evidence synthesis findings: Introduction to the series. *Implementation Science*, 13(Suppl. 1), Article 2. doi:10.1186/s13012-017-0688-3
- Lewin, S., Glenton, C., Munthe-Kaas, H., Carlsen, B., Colvin, C. J., Gülmezoglu, M., . . . Rashidian, A. (2015). Using qualitative evidence in decision making for health and social interventions: An approach to assess confidence in findings from qualitative evidence syntheses (GRADE-CERQual). *PLoS Medicine*, 12(10), e1001895. doi:10.1371/journal.pmed.1001895
- Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: Methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, 13, 179–187. doi:10.1097/XEB.0000000000000062
- Long, A. F., & Godfrey, M. (2004). An evaluation tool to assess the quality of qualitative research studies. *International Journal of Social Research Methodology*, 7, 181–196. doi:10.1080/1364557032000045302
- Majid, U., Kandasamy, S., Arora, N., & Vanstone, M. (in press). Patient perspectives and experiences. In *HPV testing for primary cervical cancer screening: A health technology assessment*. Ottawa: Canadian Agency for Drugs and Technologies in Health.
- Masood, M., Thaliath, E. T., Bower, E. J., & Newton, J. T. (2011). An appraisal of the quality of published qualitative dental research. *Community Dentistry and Oral Epidemiology*, 39, 193–203. doi:10.1111/j.1600-0528.2010.00584.x
- McInnes, E., & Wimpenny, P. (2008). Using qualitative assessment and review instrument software to synthesise studies on older people's views and experiences of falls prevention. *International Journal of Evidence-Based Healthcare*, 6, 337–344. doi:10.1111/j.1744-1609.2008.00104.x
- Melia, K. M. (2010). Recognizing quality in qualitative research. In *SAGE handbook of qualitative health research* (pp. 559–574). Thousand Oaks: Sage.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. doi:10.1371/journal.pmed.1000097
- Noblit, G. W., Hare, R. D., & Hare, R. D. (1988). *Meta-ethnography: Synthesizing qualitative studies Particularities: Collected Essay on Ethnography and Education 44*, 93-123. USA: Sage Publications.
- Noyes, J., & Popay, J. (2007). Directly observed therapy and tuberculosis: How can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis. *Journal of Advanced Nursing*, 57, 227–243. doi:10.1111/j.1365-2648.2006.04092.x
- O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, 89, 1245–1251. doi:10.1097/ACM.0000000000000388
- 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, 341–351. doi:10.1177/104973239800800305
- Sandelowski, M., & Barroso, J. (2002). Finding the findings in qualitative studies. *Journal of nursing scholarship*, 34(3), 213-219. <https://doi.org/10.1111/j.1547-5069.2002.00213.x>
- Sandelowski, M., & Barroso, J. (2003). Creating metasummaries of qualitative findings. *Nursing research*, 52(4), 226-233.
- Sandelowski, M., Barroso, J., & Voils, C. I. (2007). Using qualitative metasummary to synthesize qualitative and quantitative descriptive findings. *Research in nursing & health*, 30(1), 99-111. <https://doi.org/10.1002/nur.20176>

- Saini, M., & Shlonsky, A. (2012). *Systematic synthesis of qualitative research*. USA: Oxford University Press.
- Sandelowski, M. (2000). Focus on research methods whatever happened to qualitative description? *Research in Nursing & Health, 23*, 334–340.
- Sandelowski, M., & Barroso, J. (2006). *Handbook for synthesizing qualitative research*. New York, NY, USA: Springer.
- Santiago-Delefosse, M., Gavin, A., Bruchez, C., Roux, P., & Stephen, S. L. (2016). Quality of qualitative research in the health sciences: Analysis of the common criteria present in 58 assessment guidelines by expert users. *Social Science & Medicine, 148*, 142–151. doi:10.1016/j.socscimed.2015.11.007
- Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in qualitative evaluation: A framework for assessing research evidence*. London: Cabinet Office, Government Chief Social Researcher's Office.
- Stige, B., Malterud, K., & Midtgarden, T. (2009). Toward an agenda for evaluation of qualitative research. *Qualitative Health Research, 19*, 1504–1516. doi:10.1177/1049732309348501
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), Article 45. doi:10.1186/1471-2288-8-45
- Thorne, S. (2017). Metasynthetic madness: What kind of monster have we created? *Qualitative Health Research, 27*, 3–12. doi:10.1177/1049732316679370
- Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated Criteria for Reporting Qualitative Research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care, 19*, 349–357. doi:10.1093/intqhc/mzm042
- Vanstone, M., Kandasamy, S., Giacomini, M., DeJean, D., & McDonald, S. D. (2016). Pregnant women's perceptions of gestational weight gain: A systematic review and meta-synthesis of qualitative research. *Maternal & Child Nutrition, 13*, e12374. doi:10.1111/mcn.12374
- Walsh, D., & Downe, S. (2006). Appraising the quality of qualitative research. *Midwifery, 22*, 108–119. doi:10.1016/j.midw.2005.05.004
- Whiting, P., Wolff, R., Mallett, S., Simera, I., & Savović, J. (2017). A proposed framework for developing quality assessment tools. *Systematic Reviews, 6*(1), Article 204. doi:10.1186/s13643-017-0604-6
- Yardley, L. (2008). Demonstrating validity in qualitative psychology. In *Qualitative psychology: A practical guide to research methods*. 2, 235–251.

Author Biographies

Umair Majid was a MSc Candidate in the Health Research Methodology Program at McMaster University at the time this work was conducted. He is currently a PhD student at the Institute of Health Policy, Management, and Evaluation, University of Toronto, Ontario, Canada

Meredith Vanstone is an Assistant Professor in the Department of Family Medicine, member of the Centre for Health Economics and Policy Analysis, and Scientist in the McMaster Faculty of Health Sciences program for Education Research, Innovation & Theory at McMaster University, Hamilton, Ontario, Canada