Critical appraisal of qualitative research: necessity, partialities and the issue of bias

Veronika Williams, Anne-Marie Boylan, David Nunan

Introduction
Qualitative evidence allows researchers to analyse human experience and provides useful exploratory insights into experiential matters and meaning, often explaining the ‘how’ and ‘why’. As we have argued previously, qualitative research has an important place within evidence-based healthcare, contributing to among other things policy on patient safety, prescribing, and understanding chronic illness. Equally, it offers additional insight into quantitative studies, explaining contextual factors surrounding a successful intervention or why an intervention might have ‘failed’ or ‘succeeded’ where effect sizes cannot. It is for these reasons that the MRC strongly recommends including qualitative evaluations when developing and evaluating complex interventions.

‘Qualitative research’ is an umbrella term that refers to various research methodologies including ethnography, phenomenology, action research, discourse analysis and grounded theory that use textual and/or visual rather than numerical data to explore meaning and perspectives of phenomena. However, it is often considered one unified practice is difficult to justify. That being said, simply to omit a crucial element of evidence-based practice is difficult to justify. That being said, simply to omit a crucial element of evidence-based healthcare, then any argument to omit a crucial element of evidence-based healthcare is now increasingly widely supported (discussed in paper 1), the role of appraising the quality of qualitative health research is still debated. Despite a large body of literature focusing on appraisal and rigour, often referred to as ‘trustworthiness’ in qualitative research, there remains debate about how to—and even whether to—critically appraise qualitative research. However, if we are to make a case for qualitative research as integral to evidence-based healthcare, then any argument to omit a crucial element of evidence-based practice is difficult to justify. That being said, simply applying the standards of rigour used to appraise studies based on the positivist paradigm (Positivism depends on quantifiable observations to test hypotheses and assumes that the researcher is independent of the study. Research situated within a positivist paradigm is based purely on facts and consider the world to be external and objective and is concerned with validity, reliability and generalisability as measures of rigour, would be misplaced given the different epistemological underpinnings of the two types of data.

Given its scope and its place within health research, the robust and systematic appraisal of qualitative research to assess its trustworthiness is as paramount to its implementation in clinical practice as any other type of research. It is important to appraise different qualitative studies in relation to the specific methodology used because the methodological approach is linked to the ‘outcome’ of the research (eg, theory development, phenomenological understandings and credibility of findings). Moreover, appraisal needs to go beyond merely describing the specific details of the methods used (eg, how data were collected and analysed), with additional focus needed on the overarching research design and its appropriateness in accordance with the study remit and objectives.

Poorly conducted qualitative research has been described as ‘worthless, becomes fiction and loses its utility’. However, without a deep understanding of concepts of quality in qualitative research or at least an appropriate means to assess its quality, good qualitative research also risks being dismissed, particularly in the context of evidence-based healthcare where end users may not be well versed in this paradigm.

Critical appraisal of qualitative research
Is it necessary?
Although the importance of qualitative research to improve health services and care is now increasingly widely supported (discussed in paper 1), the role of appraising the quality of qualitative health research is still debated. Despite a large body of literature focusing on appraisal and rigour, often referred to as ‘trustworthiness’ in qualitative research, there remains debate about how to—and even whether to—critically appraise qualitative research. However, if we are to make a case for qualitative research as integral to evidence-based healthcare, then any argument to omit a crucial element of evidence-based practice is difficult to justify. That being said, simply applying the standards of rigour used to appraise studies based on the positivist paradigm (Positivism depends on quantifiable observations to test hypotheses and assumes that the researcher is independent of the study. Research situated within a positivist paradigm is based purely on facts and consider the world to be external and objective and is concerned with validity, reliability and generalisability as measures of rigour, would be misplaced given the different epistemological underpinnings of the two types of data.

How is appraisal currently performed?
Appraising the quality of qualitative research is not a new concept—there are a number of published appraisal tools, frameworks and checklists in existence. An important and often overlooked point is the confusion between tools designed for appraising methodological quality and reporting guidelines designed to assess the quality of methods reporting. An example is the Consolidate Criteria for Reporting Qualitative Research (COREQ), which was designed to provide standards for authors when reporting qualitative research but is often mistaken for a methods appraisal tool.

Broadly speaking there are two types of critical appraisal approaches for qualitative research: checklists and frameworks. Checklists have often been criticised for confusing quality in qualitative research with ‘technical fixes’, resulting in the erroneous prioritisation of particular aspects of methodological processes over others (eg, multiple coding and triangulation). It could be argued that a checklist approach adopts the positivist paradigm, where the focus is on objectively assessing ‘quality’ where the assumptions


© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

Correspondence to: Dr Veronika Williams, Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford OX2 6GG, UK; veronika.williams@phc.ox.ac.uk

Nuffield Department of Primary Care Health Sciences, University of Oxford, Radcliffe Observatory Quarter, Oxford, UK

10.1136/bmjebm-2018-111132
is that the researcher is independent of the research conducted. This may result in the application of quantitative understandings of bias in order to judge aspects of recruitment, sampling, data collection and analysis in qualitative research papers. One of the most widely used appraisal tools is the Critical Appraisal Skills Programme (CASP) and along with the JBI QARI (Joanna Briggs Institute Qualitative Assessment and Assessment Instrument) presents examples which tend to mimic the quantitative approach to appraisal. The CASP qualitative tool follows that of other CASP appraisal tools for quantitative research designs developed in the 1990s. The similarities are therefore unsurprising given the status of qualitative research at that time.

Frameworks focus on the overarching concepts of quality in qualitative research, including transparency, reflexivity, dependability and transferability (see Box 1). However, unless the reader is familiar with these concepts—their meaning and impact, and how to interpret them—they will have difficulty applying them when critically appraising a paper.

The main issue concerning currently available checklist and framework appraisal methods is that they take a broad brush approach to ‘qualitative’ research as whole, with few, if any, sufficiently differentiating between the different methodological approaches (eg, Grounded Theory, Interpretative Phenomenology, Discourse Analysis) nor different methods of data collection (interviewing, focus groups and observations). In this sense, it is akin to taking the entire field of ‘quantitative’ study designs and applying a single method or tool for their quality appraisal. In the case of qualitative research, checklists, therefore, offer only an incomplete understanding of good ‘quality’ in qualitative research. Likewise, current framework methods do not take into account how concepts differ in their application across the variety of qualitative approaches and, like checklists, they also do not differentiate between different qualitative methodologies.

Box 1 Concepts of rigour or trustworthiness within qualitative research

Transferability: the extent to which the presented study allows readers to make connections between the study’s data and wider community settings, ie, transfer conceptual findings to other contexts.

Credibility: extent to which a research account is believable and appropriate, particularly in relation to the stories told by participants and the interpretations made by the researcher.

Reflexivity: refers to the researchers’ engagement of continuous examination and explanation of how they have influenced a research project from choosing a research question to sampling, data collection, analysis and interpretation of data.

 Transparency: making explicit the whole research process from sampling strategies, data collection to analysis. The rationale for decisions made is as important as the decisions themselves.

It is important to note that we are not referring to these prejudices as ‘biases’, as to do so is to invoke quantitative understandings of bias which are not wholly compatible with qualitative research. To avoid confusion, we will refer to them as partialities. These partialities do not relate to reliability, validity and reproducibility (the quantitative concepts of rigour) but rather to transferability, credibility, reflexivity and dependability (see Box 1).

However, we often talk about these concepts in general terms, and it might be helpful to give some explicit examples of how the ‘technical processes’ affect these, for example, partialities related to:

- Selection: recruiting participants via gatekeepers, such as healthcare professionals or clinicians, who may select them based on whether they believe them to be ‘good’ participants for interviews/focus groups.
- Data collection: poor interview guide with closed questions which encourage yes/no answers and leading questions.
- Reflexivity and transparency: where researchers may focus their analysis on preconceived ideas rather than ground their analysis in the data and do not reflect on the impact of this in a transparent way.

The lack of tailored, method-specific appraisal tools has potentially contributed to the poor uptake and use of qualitative research for informing evidence-based decision making. To improve this situation, we propose the need for more robust quality appraisal tools that explicitly encompass both the core design aspects of all qualitative research (sampling/data collection/analysis) but also considered the specific partialities that can be presented with different methodological approaches. Such tools might draw on the strengths of current frameworks and checklists while providing users with sufficient understanding of concepts of rigour in relation to the different types of qualitative methods. We provide an outline of such tools in the third and final paper in this series.

Conclusion

As qualitative research becomes ever more embedded in health science research, and in order for that research to have better impact on healthcare decisions, we need to rethink critical appraisal and develop tools that allow differentiated evaluations of the myriad of qualitative methodological approaches rather than continuing to treat qualitative research as a single unified approach.

Correction notice This article has been updated since its original publication to include a new reference (reference 1.)

Contributors VW and DN: conceived the idea for this article. VW: wrote the first draft. AMB and DN: contributed to the final draft. All authors approved the submitted article.

Competing interests None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Veronika Williams http://orcid.org/0000-0001-5660-8224
David Nunan http://orcid.org/0000-0003-4597-1276

References
5 Bury M. Chronic illness as biographical disruption. *Social Health Illn* 1982;4:167–82.
8 Smith B, McGannon KR. Developing rigor in qualitative research: problems and opportunities within sport and exercise psychology. *Int Rev Sport Exerc Psychol* 2018;11:301–21.
17 Garside R. Should we appraise the quality of qualitative research reports for systematic reviews, and if so, how? *Innovation: The European Journal of Social Science Research* 2014;27:67–79.
22 Harris J. Supplementary guidance for inclusion of qualitative research in Cochrane systematic reviews of interventions. 2011.
26 CASP [Critical Appraisal Skills Programme], date unknown. [http://www.phru.nhs.uk/Pages/PHD/CASP.htm](http://www.phru.nhs.uk/Pages/PHD/CASP.htm).