
Article:
Munthe-Kaas, H.M. orcid.org/0000-0002-6051-9010, Booth, A. orcid.org/0000-0003-4808-3880, Sommer, I. orcid.org/0000-0003-3592-1507 et al. (4 more authors) (2024) Developing CAMELOT for assessing methodological limitations of qualitative research for inclusion in qualitative evidence syntheses. Cochrane Evidence Synthesis and Methods, 2 (6). e12058. ISSN 2832-9023

https://doi.org/10.1002/cesm.12058

Received: 26 January 2024 | Accepted: 17 April 2024

DOI: 10.1002/cesm.12058

Cochrane Evidence Synthesis and Methods

METHODS ARTICLE

Developing CAMELOT for assessing methodological limitations of qualitative research for inclusion in qualitative evidence syntheses

Heather M. Munthe-Kaas[1] | Andrew Booth[2] | Isolde Sommer[3] | Sara Cooper[4] | Ruth Garside[5] | Karin Hannes[6] | Jane Noyes[7] | The CAMELOT Development Group

[1]Centre for Epidemic Interventions Research, Norwegian Institute of Public Health, Oslo, Norway

[2]Sheffield Centre for Health & Related Research (SCHARR), University of Sheffield, Sheffield, UK

[3]Department for Evidence-based Medicine and Evaluation, University for Continuing Education Krems, Krems, Austria

[4]Cochrane South Africa, South African Medical Research Council, Cape Town, South Africa

[5]University of Exeter, Exeter, UK

[6]Research Group SoMeTHin'K, Faculty of Social Sciences, KU Leuven, Leuven, Belgium

[7]School of Medical and Health Sciences, Bangor University, Bangor, UK

Correspondence

Heather M. Munthe-Kaas, Centre for Epidemic Interventions Research, Norwegian Institute of Public Health, Oslo, Norway.
Email: heather.munthe-kaas@fhi.no

Funding information

Cochrane Methods Innovation Fund, Grant/Award Number: 1328523

Abstract

Introduction: Qualitative evidence is increasingly incorporated into decision-making processes. Assessing the methodological limitations of primary studies is critical to making an overall assessment of confidence in findings from qualitative evidence syntheses (QES) using GRADE-CERQual. Current critical appraisal tools were not developed specifically for use in Cochrane reviews or GRADE-CERQual, and few are evidence-based. The aim of CochrAne qualitative Methodological LimitatiOns Tool (CAMELOT) was to address this gap.

Methods: We undertook this project in four stages: (1) systematic literature search to identify existing tools, (2) identification of evidence to support inclusion of potential CAMELOT domains (3) consensus survey to agree on the inclusion and definition of CAMELOT domains, and (4) human-centered design approach to develop and refine CAMELOT by exploring user experience.

Results: CAMELOT is a new evidence-based tool for assessing the methodological strengths and limitations of primary qualitative research studies in a QES. CAMELOT is comprised of 12 domains: four *Meta domains* that encourage review authors to consider those characteristics of the primary study that are beyond how the study was carried out, but which inform the conduct and design of the study, and eight *Method domains* which encourage review authors to consider how the study was designed, planned and/or conducted, and how study conduct and design fits with the information provided in the four meta domains. Review authors make an assessment by identifying any concerns regarding the methods used in the study and considering the appropriateness of fit between the *Meta* and *Method domains*.

Conclusion: CAMELOT provides review authors with a transparent and systematic method to assess methodological limitations of primary qualitative studies. CAMELOT incorporates qualitative principles and focuses on appropriateness of fit between *Meta* and *Method domains*. In line with iterative tool development

approach, CAMELOT will continue to be revised over time following extensive user testing and piloting.

KEYWORDS
consensus development, critical appraisal, Delphi, GRADE-CERQual, methodological limitations, qualitative evidence synthesis, qualitative research

1 | INTRODUCTION

This paper reports the development of CochrAne qualitative Methodological LimitatiOns Tool (CAMELOT), a new tool for assessing methodological limitations in primary qualitative studies.

Qualitative evidence syntheses (QESs)—also called systematic reviews of qualitative evidence—are increasingly recognized as an important way to incorporate qualitative research into decision-making processes, including guideline development and policy formulation [1–3]. Assessing the methodological limitations of included studies is key to the overall assessment of evidence from QESs. The methodological limitations of included primary studies and the extracted data play a critical role in determining how much confidence to place in synthesized review findings. The "Confidence in the Evidence from Reviews of Qualitative Research" (GRADE-CERQual) approach was developed to support the use of findings from QES in decision-making [4]. The approach aims to transparently and systematically assess how much confidence to place in findings from QESs. Application of GRADE-CERQual is an expectation in Cochrane QESs and since its release has been used globally in over 233 QESs [5]. CAMELOT has been designed with GRADE-CERQual in mind, but can also be used when review authors do not use GRADE-CERQual. GRADE-CERQual bases its assessment of confidence on four components: the methodological limitations of individual studies contributing to a review finding, the adequacy of data supporting a review finding, the relevance of the studies contributing to a finding from which data has been extracted, and the coherence of each review finding [6]. These four components are considered together when assessing how much confidence to place in synthesized findings from QESs.

1.1 | Current practice in tool selection

The Cochrane Qualitative and Implementation Methods Group's (QIMG) handbook chapter 21 on QES outlines a set of minimum criteria when selecting a tool, although no tool offers a complete fit in supporting the assessment of the GRADE-CERQual methodological limitations component [7]. Although not encouraged to do so, as a default, authors of Cochrane QESs either adapt an existing appraisal tool, [8] use a familiar tool [9] or pick a tool from a list [10]. It is, however, widely accepted that tools specifically intended for qualitative research should be used as criteria for rigor used within quantitative research would be inappropriate [11].

1.2 | Why was a new tool needed?

Review authors' assessment of study's methodological strengths and limitations (rigor) is a process referred to as critical appraisal and "seeks to assess the validity and reliability of a primary research study and its findings" [12]. This process has, to date, posed significant challenges to review authors when conducting QESs [13, 14]. Review authors currently face a choice of more than 100 critical appraisal tools, [15, 16] few of which report to be evidence-based. Furthermore, most of these tools were developed before the inclusion of patient and public involvement as an essential element of qualitative research [13]. These tools vary greatly in their aims –some include domains of reporting quality, while others seek to identify the technical quality or how well a study was conducted.

Many current tools support end users when assessing "research evidence to judge its trustworthiness, its value and relevance in a particular context." [17] Such tools seek to evaluate the quality or utility of the study and make an overall assessment using a narrative description (e.g., "good/bad" judgments). Some tools even attempt to quantify the overall assessment using a summary score (e.g., numerical scales). However, assessing methodological limitations within the context of QES where evidence may be used to inform practice or policy may differ from assessments conducted as a teaching exercise, for example. Assessing the overall quality of a single qualitative study is categorically different from assessing a group of studies that contribute data to a synthesized review finding [18]. Assessing a group of studies requires looking past the quality of individual studies to focus on methodological limitations across the group of studies (body of evidence) that could weaken trustworthiness of a review finding.

Given the key role of a methodological limitations assessment in informing an overall assessment of confidence in review findings, it is important to have an evidence-based tool designed for use specifically within the context of QES and GRADE-CERQual.

2 | OBJECTIVE

We aimed to identify or develop an evidence-informed approach called CAMELOT to assess the methodological limitations of primary qualitative research within the context of GRADE-CERQual assessments. Specifically, we sought to address the following research questions:

1. What evidence is there that criteria included in current critical appraisal tools are good indicators of trustworthiness or rigor of qualitative research?
2. Are there gaps in existing criteria for assessing methodological limitations?
3. What do methodological experts and relevant stakeholders agree are the key domains to consider when assessing the methodological strengths and limitations of primary qualitative research within the context of QES and GRADE-CERQual?
4. What are users' experiences with CMELOT to assessing methodological strengths and limitations of primary qualitative research?

3 | MATERIALS AND METHODS

We followed a rigorous approach and followed good practice guidelines related to tool development and Delphi surveys [19]. We addressed the above research questions in four stages:

- Phase 1: Identification of current critical appraisal tools and synthesis of included criteria [13].
- Phase 2: Identification of evidence related to how current checklist criteria influence trustworthiness of qualitative research, identification of gaps (criteria not currently covered), and development of domains for assessing methodological strengths and limitations.
- Phase 3: Consensus on CAMELOT domains and guidance.
- Phase 4: User experience with CAMELOT.

3.1 | Phase 1. Identification of checklists and synthesis of included criteria

Methods for this stage, a systematic search and analysis of items from available checklists, are reported in a previously published article [13].

3.2 | Phase 2. Identification of evidence for criteria for assessing methodological strengths and limitations of primary qualitative research

The findings from Phase 1 were used to inform the research agenda for Phase 2 which was led by a working group of seven researchers from diverse and relevant backgrounds (e.g., nursing, information science, social welfare, public health, medical sociology), fields of research, qualitative traditions and methodological expertise with an interest in improving methods for considering methodological strengths and limitations of primary qualitative research (most of whom are also co-founders and co-ordinators of the GRADE-CERQual approach). Their positionality offers an advantageous perspective on what review teams look for when conducting critical appraisal of primary studies within QESs. Three researchers sit in the Cochrane Qualitative and Implementation Methods Group (Jane Noyes, Andrew Booth, and Ruth Garside) and four (Jane Noyes, Andrew Booth, Isolde Sommer, and Karin Hannes) have experience with Delphi processes. The research team occupies various roles in relation to national and international evidence synthesis and guideline production activities.

In conducting this research, we acknowledge the importance of considering equity, diversity, and inclusion. We recognize that the findings from this study differentially impact communities of end users and researchers. We sought to ensure the representation of a diverse set of stakeholders by recruiting panelists from different backgrounds related to geographical location, type of institution and role, experience level, and type of research. We encouraged diverse representation of gender and age. Despite best efforts, we recognize that data collection, analysis, and interpretation of findings may be shaped by our own experiences and/or epistemological stance.

Between 2017 and 2023 the research team held a series of meetings to discuss each of the criteria identified in Phase 1. We systematically reviewed the list of criteria from Phase 1 [13] and (a) excluded those criteria that were explicitly related to reporting quality, the "value" of the

research, or the degree to which the study was qualitative and whether qualitative approach was appropriate, and (b) synthesized criteria that were related.

We evaluated the relevance of each criterion for assessing methodological strengths and limitations of primary qualitative research within the context of qualitative evidence synthesis. We examined published evidence to support inclusion of that criterion within a critical appraisal tool (i.e., do concerns regarding component A influence the trustworthiness of findings from a qualitative study?). In the context of this study, evidence encompasses any empirical evidence, theory or framework, or observations (anecdotal evidence) that explore the influence of a specific criterion on the trustworthiness of qualitative research findings. For each of the criteria we examined, Andrew Booth developed and carried out a search. A researcher (Heather M. Munthe-Kaas, Andrew Booth, Isolde Sommer, Sara Cooper, Jane Noyes, Ruth Garside, and Karin Hannes) reviewed the identified references and shared the findings with the team before an online discussion focused on making a preliminary recommendation regarding inclusion of the criteria, sometimes rewording/reframing the criterion. The results of each discussion were posted in an open-access repository [20].

3.3 | Phase 3: Consensus on CAMELOT domains and guidance

The results of Phase 2 were summarized and used as the basis for a consensus survey, in accordance with good practice standards for methodological guidance and tool development [19] and guidance for conducting and reporting Delphi studies [21]. To avoid excessive focus on the specific wording of each criterion, and to harmonize with analogous contemporary approaches (e.g., Risk of Bias tool development [22]), the survey was operationalized around corresponding domains (e.g., Researchers).

An online Delphi survey (eDelphi) was initiated in June 2023 with the aim of achieving consensus on the list of proposed domains resulting from Phase 2. The Delphi technique is commonly used to gather opinions and feedback from relevant experts, and to achieve consensus on an output [23]. The eDelphi survey consisted of three stages:

1. Introductory seminar.
2. eDelphi survey—two rounds of the eDelphi process were used to exclude irrelevant criteria, and compile and prioritize relevant criteria and their definitions. Originally planned to include three rounds, the very high consensus from the first round made a third round unnecessary.
3. Consensus meeting—an online workshop to finalize a list of CAMELOT domains.

We aimed to include 56 methodological experts and 49 other stakeholders with experience produing or using qualitative evidence. See Appendix S1 and the published protocol [24] for more details regarding methods used in the eDelphi survey and consensus webinar.

3.4 | Phase 4. User experience

We used human-centered design methods [25] to develop and refine CAMELOT, including instructions and examples (Figure 1). Human-centered design methods are intended to ensure that end users are at the center of a process where the goal is to develop a user-friendly and clear tool that sufficiently meets the user's needs.

3.4.1 | Sampling and recruitment

We used convenience and snowballing methods to identify and recruit user testing participants. We aimed to recruit 5–10 participants from diverse backgrounds in terms of level of experience, academic discipline, and native language. Participants were invited via email to participate and provided with information regarding the goal of the user test. We obtained informed consent from those who agreed to participate.

3.4.2 | Data collection

We conducted six online semistructured interviews (conversation guide adapted from Morville's honeycomb model [26–28] and available upon request) using the think-aloud method and focused on usability and understandability to gather feedback from review teams and refine the

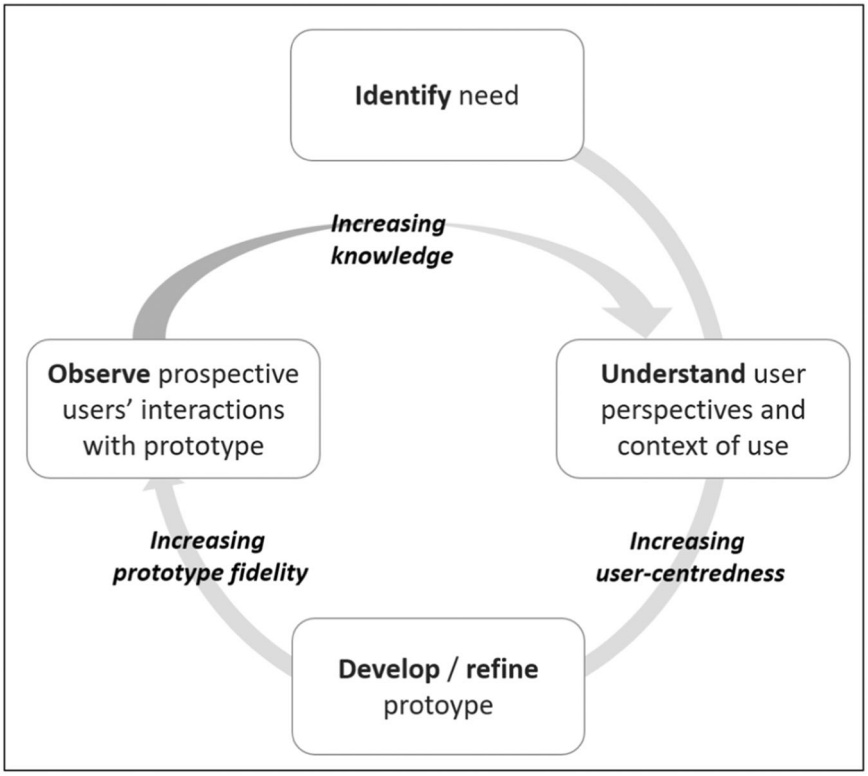

FIGURE 1 Human-centered design (adapted from Munthe-Kaas[18]).

tool [29, 30]. We recorded and transcribed the interviews using Microsoft Teams. Transcripts were anonymized, coded, and stored in a secure folder.

3.4.3 | Data analysis

We used a framework adapted from Rosenbaum 2010 [31] to categorize and analyze the seriousness of user experience issues using the following categories: showstopper (major changes needed), consider changes, minor problem, positive feedback but consider changes, positive feedback with no changes, not relevant [32].

3.5 | Ethical considerations

This study did not collect sensitive information. All feedback was anonymised, and no vulnerable groups were involved. The study was therefore outside the remit of the Norwegian Regional Ethical Committee.

For the consensus survey, informed consent was elicited when panelists submitted an initial Google form indicating their wish to participate. We obtained consent from panelists before recording the introductory and consensus webinars. At the introductory seminar panelists agreed how intellectual contribution would be recognized in ensuing publications or project outputs.

For the user testing, we obtained informed consent and anonymized all data, however, no Data Protection Impact Assessment approval was required.

4 | RESULTS

We present two main categories of results: findings related to Phases 2, 3, and 4 of the development process, and a description of the subsequent CAMELOT approach. Findings from Phase 1 are published elsewhere [24].

4.1 | Findings from Phase 2: Identification of evidence for criteria for assessing methodological strengths and limitations of primary qualitative research

Phase 2 generated 12 domains that were determined to impact on the trustworthiness of findings from primary qualitative research and thus considered potentially important for making assessments of methodological strengths and limitations. The working group agreed an interim definition for each of the 12 domains, based on findings from the literature search. A summary of the findings from empirical, theoretical, or anecdotal evidence related to the 17 relevant criteria that were promoted from Phase 1 is summarized and available on an open-access repository [20].

4.2 | Findings from Phase 3: Consensus on CAMELOT domains and guidance

Sixty invitees agreed to participate in the Delphi and were sent a link to the eDelphi survey. The 12 domains and definitions from Phase 2 were presented to panelists in Round 1 of the eDelphi (see Appendix S2).

A total of 41–42 participants responded to Round 1 of the Delphi survey (number of responses varied by question). More than 60% of panelists indicated that each of the 12 domains was important or very important when making an assessment of methodological strengths and limitations of primary qualitative research (see Appendix S2). Panelists suggested eight additional domains (researcher reflexivity, equity, funding, complete reporting, prospective registering of protocols, sampling, description of methodology, and fit between components). Most suggestions had been previously identified within existing domains. However, two of the suggested domains had not been addressed and were added for consideration in Round 2: Equity was added under "ethical and equity considerations," and completeness of reporting was added to determine its perceived importance when making assessments of methodological strengths and limitations.

4.2.1 | Round 2

In Round 2, 35–37 panelists responded and demonstrated a strong consensus in relation to satisfaction with definitions for domains promoted from Round 1 (see Appendix S3). More than 50% of panelists were satisfied or very satisfied with the current definitions (see Appendix S3). However, feedback from panelists suggested merging two domains, specifically *Research question* and *Research aim*. Additionally, numerous respondents suggested dividing *ethics and equity considerations* into two separate domains: *ethical considerations* and *equity, diversity and inclusion considerations*. Panelists also provided comments on specific words and terminology. These were explored using content analysis and revised as appropriate. Sixty percent of panelists ($N = 21/35$) agreed that *completeness of reporting* was important or very important when considering methodological strengths and limitations.

The working group further revised definitions to be presented to panelists in the consensus webinar.

4.2.2 | Consensus webinar

We achieved 100% consensus for each of the 12 domains and definitions presented. The webinar concluded by presenting the overall CAMELOT approach including brief instructions for how to consider each domain and assess appropriateness between domains. Panelists' first impressions were largely positive, however, some voiced concerns that the approach may be considered long, resource intensive, or overwhelming for some end users. Panelists generally agreed that the approach seemed compatible with different types of qualitative research, as well as with differing experience or competence among review teams or end users (e.g., experienced researchers may truncate the process, whereas less experienced researchers may follow the prompts more systematically).

4.3 | Findings from Phase 4: User experience

We conducted semistructured interviews with six participants to explore user experience of CAMELOT which led to iterative revision based on the positive and negative feedback. The participants had a range of experience related to qualitative research and QES (from no experience conducting a QES, to having led many QESs). Several of the participants did not have English as their native language and were asked specifically to consider understandability of terminology. The first two participants gave feedback on the CAMELOT figure, instructions, and tables. The third participant attempted to apply CAMELOT during the user test and the researcher identified a showstopper, namely that there were too many assessments of fit to make (at that time numbering 72). The CAMELOT instructions and table were substantially revised and presented to

two more participants for feedback (resulting in revisions to table format). The last three participants applied CAMELOT alone and provided written feedback on their experience.

4.3.1 | Positive feedback

Participants were pleased to see the inclusion and stronger emphasis of domains such as theory and equity, diversity, and inclusion which are absent in some of the more commonly used tools. They provided suggestions regarding language and format. All participants were pleased with the tips and examples sections and requested sufficient examples to illustrate the diversity of primary studies and methodological limitations that could be identified. The participants' general impression was that CAMELOT seemed more complex and time consuming than currently commonly used tools, but that its comprehensiveness would challenge review authors to consider many aspects related to methodological limitations.

4.3.2 | Constructive feedback and corresponding changes

Many participants described being confused and overwhelmed by the amount of information included in the figure illustrating CAMELOT (an earlier version of Figure 1). Most participants needed several minutes to look at it and read through it multiple times. It was revised iteratively based on this feedback to the current version (Figure 1).

Participants were confused by the instructions for making a CAMELOT assessment and unclear of how they would work in practice. Participants suggested changes to language and format in the instructions.

Table 1 below illustrates which criteria/domains were included/not promoted during each phase of this project.

4.4 | CAMELOT

CAMELOT is an evidence-based tool for making an assessment of the methodological strengths and limitations of primary qualitative research studies that are included in a QES. It is primarily intended for reviews authors conducting QESs, but is also useful for other end users who need to assess the methodological limitations of a primary study. Review authors make an assessment by identifying any concerns regarding the methods used in the study and considering the appropriateness of the fit between the Meta and Method domains.

CAMELOT aims to better reflect the principles of qualitative research when assessing the methodological strengths and limitations of qualitative research. This qualitative approach to assessment seeks to encourage reflective consideration and transparency by focusing on and identifying concerns with appropriateness or fit between CAMELOT methodological domains. The term *appropriateness of fit* in CAMELOT refers to the degree to which the design and conduct of the study is: (1) well suited to address the research aim and/or question(s), (2) aligns with the researchers and/or stakeholders involved in the study, and/or; (3) is suitable given the study context. As with assessments made using any critical appraisal tool, assessments of appropriateness of fit are subjective judgments dependent on the review authors' experience and background. The assessments may be different between authors or review teams, reiterating the importance of undertaking the process using consensus methods by the review authors and to include explanations to ensure transparency.

CAMELOT does not entail the calculation of summary scores or overall assessments such as high, medium or low quality. Instead, it supports review authors in making assessments of the appropriateness of fit between Method and Meta domains to arrive at an overall assessment of level of concerns regarding methodological limitations, expressed as no or minimal concerns, minor concerns, moderate concerns or serious concerns (Figure 2).

4.5 | CAMELOT domains

There are 12 CAMELOT domains, divided into Meta domains and Method domains (see description below). Table 2 provides an overview of the definitions of each domain. Instructions and tips for each domain are further detailed https://zenodo.org/records/10973143.

4.6 | Overview of CAMELOT instructions

In the detailed description referred to above, we provide review authors with items to consider when extracting or coding data for each domain, and to note concerns and where important information is missing. We also provide tips for where to find the information in a study report, as

TABLE 1 Inclusion of criteria/domains during each phase.

Criteria identified from systematic review	Criteria included in Phase 2	Domains included in Phase 3	Domains included in Phase 4	CAMELOT domains
Is there a clear statement of findings?	Is there a clear statement of findings?	Presentation of findings	Presentation of findings	Presentation of findings
Was the data analysis sufficiently rigorous? Did the researchers achieve saturation?	Was the data analysis sufficiently rigorous? Did the researchers challenge their findings, for example, disconfirmatory cases?	Data analysis and interpretation	Analysis and interpretation	Analysis and interpretation
Was the data collected in a way that addressed the research issue? Did the researcher spend sufficient time in the research setting?	Was the data collection reported? This is also about reporting rather than identifying a methodological limitation? Was the data collected appropriately? Did the researcher spend sufficient time in the research setting?	Data collection	Data collection	Data collection
Who are the participants, how were they selected and were the methods for selection appropriate?	Were the methods for participation selection appropriate?	Participant recruitment and selection	Participant recruitment and selection	Participant recruitment and selection
Has the research team considered their role in the research process and any influence it may have on the research process or findings? Have authors discussed/assessed the overall rigor of the research study including strengths and limitations of the research?	Has the research team considered their role in the research process and any influence it may have on the research process or findings and how any funding may have impacted the study?	Researchers	Researchers	Researchers
Was there a statement of the aims of the research?	Was there an aim of statement of research	Research aim Research question	Research aim and question	Research aim and question
	Was the research design described - This is also about reporting rather than identifying a methodological limitation?	Research process	Research approach	Research strategy
Was the research design appropriate to address the aims of the research?	Was the design appropriate to address the aims? Is there coherence between the context of the study and the aim of the study?			
Have ethical issues been taken into consideration?	Are there any reasons for ethical concerns that could have influenced the findings of the study?	Ethical considerations	Ethical considerations	Ethical considerations

TABLE 1 (Continued)

Criteria identified from systematic review	Criteria included in Phase 2	Domains included in Phase 3	Domains included in Phase 4	CAMELOT domains
Did the authors include/discuss a theoretical perspective?	If the authors applied a theoretical perspective, is it appropriate? Did they apply it appropriately? Cross-reference to Cochrane guidance: what counts as theory? Different criteria for theory generation and theory utilization	Theoretical framework	Theory	Theory
Did the authors conduct a review of the literature?	Did the authors conduct a review of the literature?			
Is there an audit trail?	Is there an audit trail?			
Were end users involved in the development of the research study?	Who were the stakeholders and were they properly involved? (see Burls[17])	Stakeholders		Stakeholders
	Were end users involved in the development of the research study?			
		Context	Context	Context
			Equity, diversity, and inclusion considerations	Equity, diversity, and inclusion considerations
How valuable is the research?				
Is a qualitative method appropriate?				
Reporting criteria (including demographic features of the study)				
Is this a qualitative study?				
Did the authors consider/report practicalities of conducting project, and were they realistic?				
Are the authors credible?				
Was there disclosure of finding sources?				

Abbreviation: CAMELOT, CochrAne qualitative Methodological LimitatiOns Tool.

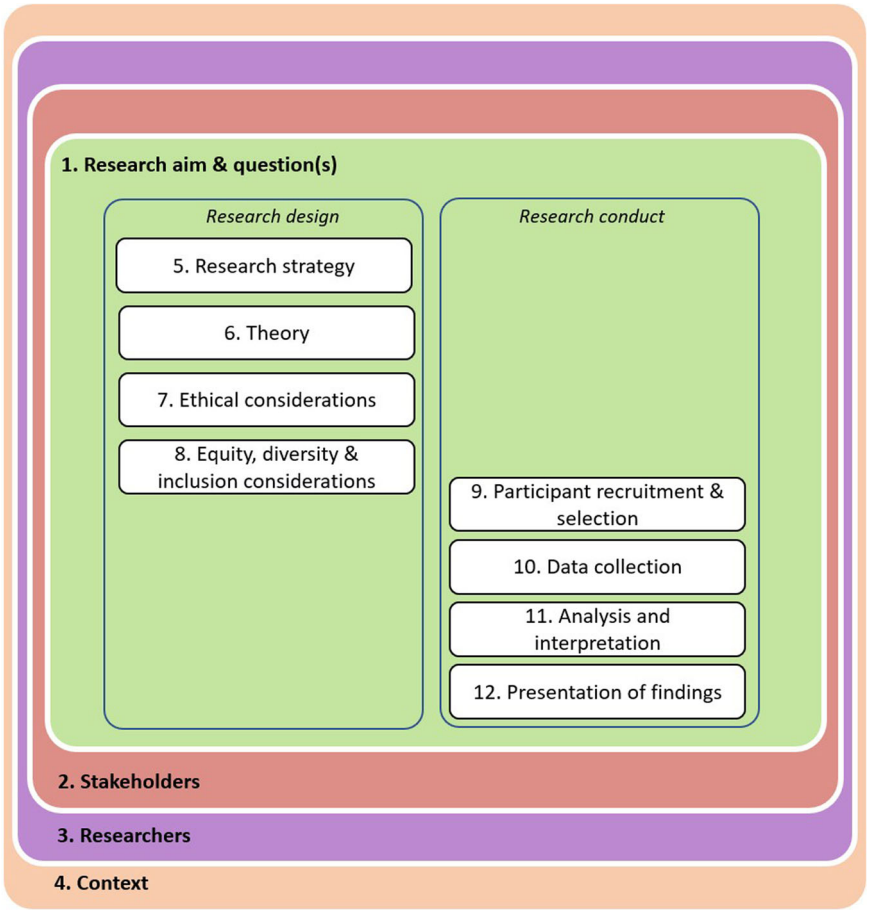

FIGURE 2 Overview of CochrAne qualitative Methodological LimitatiOns Tool (CAMELOT).

well as other notes related to terminology, specific qualitative approaches, and so on. The four Meta domains (research aim and question[s], researchers, stakeholders, context) encourage review authors to consider characteristics of the primary study beyond how the study is carried out but inform the conduct and design of the study. The eight method domains (research approach, ethical considerations, equity, diversity/inclusion considerations, theory, participant recruitment and selection, data collection, analysis and interpretation, and presentation of findings) encourage review authors to consider how the study was designed, planned, and/or conducted.

4.7 | Making a CAMELOT assessment

In Box 1, we have described an overview of the steps for making a CAMELOT assessment.

4.8 | Planning a CAMELOT assessment

When planning a CAMELOT assessment, the review team should consider incorporating the CAMELOT domains into their data extraction or coding worksheet (or software where appropriate). This may prevent duplication of work—for instance, most data extraction forms already include fields related to *research aim and question(s)*, *context*, *participant recruitment and selection*, *data collection*, *and analysis and interpretation*.

4.9 | Assessing an individual study

Whenever possible, CAMELOT should be undertaken by multiple reviewers (simultaneously or via checks for consensus). This is both considered best practice by Cochrane [33] and also provides opportunities for discussion that could potentially influence how the review team understands

TABLE 2 Definitions of CAMELOT domains.

Domain	Definition
META domains	
Research aim and question(s)	The purpose of the study and/or what questions the researchers aim to explore.
Stakeholders	Anyone with an interest (financial or otherwise) in the findings of the research study. Stakeholders are not the same as research participants in this context. Stakeholders may include, among others, funders, patient, and public participants.
Researchers	The investigators who have designed, planned, and conducted the study and their relationship to the study question, context, and/or participants.
Context	The local, national, or international setting in which the study was conducted.
METHOD domains	
Research design domains	
Research strategy	The overall intended plan, proposal, or strategy for the study. This domain refers to the overarching roadmap for carrying out the research project (also referred to as research approach, study design, or type of study). This domain does not include issues related to participant recruitment and selection, data collection and analysis, and interpretation. These are separate domains.
Ethical considerations	How the researchers considered and incorporated ethical principles and standards into decisions related to the design, planning, and conduct of the study.
Equity, diversity, and inclusion considerations	Whether and how the researchers considered: (1). Equity—including distribution of power within the research context, whether there was equitable representation and participation in the research process, particularly for underrepresented groups, the possible differential experiences or perspectives of a phenomenon of interest for different populations, and whether there was and whether unnecessary or discriminating differences in how people participate in a study (2). Diversity—including seeking out diverse experiences, perspectives, and backgrounds, inclusion of participants with diverse backgrounds, and considering how diversity can influence research findings (3). Inclusion—including the degree to which the research environment was such that all participants felt welcome and valued, whether culturally sensitive and inclusive research methods and communication strategies were employed, and whether research materials, locations, and processes were accessible for all participants.
Theory	Organization of concepts, ideas, literature, or principles into systems or frameworks that attempt to describe, explore, explain, understand, or predict a phenomenon.
Research conduct domains	
Participant recruitment and selection	How participants were identified, recruited, and selected for the research study.
Data collection	The process of gathering qualitative information (data) in the form of perspectives, experiences, or opinions from participants, and/or observations, prolonged engagement in the field by researchers to explore or answer the research questions and address the research aim.
Analysis and interpretation	The process of systematically examining, exploring, and interrogating data gathered during the study to identify themes, patterns, lines of argument and, if appropriate, theories and gain a greater understanding of the phenomenon of interest.
Presentation of findings	How the findings from the study are organized and communicated and how well they appear to represent the underpinning data.

Abbreviation: CAMELOT, CochrAne qualitative Methodological LimitatiOns Tool.

the phenomenon of interest, the development of a coding framework, how data are interpreted and the review findings are presented, and any theories that inform the primary studies or are relevant to the synthesis.

There is no gold standard, or perfect study, by which to judge studies included in a QES. Rather than focusing on how well a study was conducted and presented, concerns (or red flags) may lead review authors to pause and consider how a feature of the study design or conduct influences trust in the study findings. This thoughtful and holistic approach when making assessments contrasts with simply answering "yes" or "no" to a checklist question. For example, in a study using interviews to understand teenagers' relationships to alcohol and drugs the review team could believe that participants might view the interviewer as an authority figure and therefore withhold important perspectives or experiences. They may have no other concerns regarding the design or conduct of the study or the fit between the method and meta domains. When making

BOX 1 Making a CAMELOT assessment.

Step 1. Extract/code data

Extract or code data from the primary study related to the following domains (some of these domains will not be relevant for some studies):

Meta domains

1. Research aim and question(s)
2. Stakeholders
3. Researchers
4. Context

Method domains

Research design

5. Research strategy
6. Theory
7. Ethical considerations
8. Equity, diversity, and inclusion consideration

Research conduct

9. Participant recruitment and selection
10. Data collection
11. Analysis and interpretation
12. Presentation of findings

See below for definitions of the CAMELOT domains

Step 2. Note any comments regarding each domain. This may include problems or missing information. This step is optional but will act as an audit trail and help to inform the subsequent steps.

Step 3. Describe concerns regarding, and make assessment of, fit between domains.

- Describe concerns regarding appropriateness of fit between the Research design domains and each of the Meta domains and the Research conduct domains (as a whole). Make an assessment using the following categories to describe the fit: Excellent, Good, Fair, Poor, Unclear
- Describe concerns regarding appropriateness of fit between the Research conduct domains fit with each of the Meta domains. Make an assessment using the following categories to describe the fit: excellent, good, fair, poor, and unclear.

Step 4. Describe the level of concern regarding methodological limitations

Combine these assessments to make an overall assessment of methodological limitations by indicating the level of concern using the following categories and provide an explanation for your assessment:

- *No or minimal concerns, minor concerns, moderate concerns, serious concerns.*

Step 5. Combine assessments across studies

Combine assessments of fit across studies contributing to a review finding and indicate level of concern regarding methodological limitations using the following categories:

- *No or minimal concerns, minor concerns, moderate concerns, serious concerns.*

an assessment of the primary study, the review team may note their concerns regarding the fit between the researchers and research aim and question(s) which may lead to minor concerns regarding methodological limitations. Thus, they are not making a judgment on the overall quality of the study, but rather noting where an element of design or conduct holds potential consequences for the trustworthiness of a particular review finding.

TABLE 3 Example of using CAMELOT to inform methodological limitations assessment of a QES finding.

Review finding: Adults are skeptical about the possible connection between parental diet and offspring mental health		
Study ID	**CAMELOT assessment**	**Explanation**
Study 1 2020	Moderate concerns	Unclear researchers' relationship to participants so unclear fit between research conduct and researchers' domain.
Study 2 2019	Minor concerns	No description of ethical considerations so unclear fit between research design and research aim and question(s).
Study 3 2009	No or minimal concerns	
Study 4 2010	Serious concerns	Stakeholders and funders were from a soda and snack company and the research was not carried out independently leading to concerns about impartiality and bias. Poor fit between Research conduct and Stakeholders and Context domains.
Overall assessment of methodological limitations	Moderate concerns	Multiple and diverse concerns across studies related to research design, research conduct and fit with researchers, research aim and question(s), and stakeholders.

Abbreviations: CAMELOT, CochrAne qualitative Methodological LimitatiOns Tool; QES, qualitative evidence syntheses.

4.10 | Assessing the body of studies contributing data to a review finding

In a QES, the review team makes an assessment of methodological limitations for each review finding which will inform an overall GRADE-CERQual assessment. This overall assessment should be informed by CAMELOT assessments (no or minimal, minor, moderate, or serious concerns) for each of the individual studies that contribute data to the review finding (see Table 3 for an example).

Review findings vary considerably within a QES and the influence of concerns identified through CAMELOT may impact each review finding in different ways. For example, for review findings that are relatively uncontroversial and seemingly obvious, poor fit between the research conduct domains and the researchers' domain may not be considered important.

5 | DISCUSSION

Application of CAMELOT helps review authors and other end users to consider the methodological strengths and limitations of primary qualitative research and makes it easier to assess the trustworthiness of the findings from qualitative research. The development of CAMELOT was based on methods following good practice for tool development, [19] human-centered design, [34] and Delphi consensus surveys [35]. CAMELOT is not intended as a comprehensive checklist by which review teams look for every element under each domain. It is intended to encourage a thoughtful and thorough approach to assessing methodological strengths and limitations by incorporating the principles of qualitative research and focusing and reflecting on appropriateness of fit between different domains related to research planning, design, conduct, and the interpretation of findings where assessing methodological limitations is "utilized as part of a process of exploration and interpretation in the synthesis process" [14].

Critical appraisal is an inherently complex and time-consuming, but necessary, component of QES. User experience feedback suggests that CAMELOT may be more time-consuming and complex than other commonly used tools. However, user experience thus far has been relegated to the use of CAMELOT as a tool separate to the rest of the QES process. However, in a traditional QES data extraction form, many CAMELOT domains such as Research aim, ethical considerations, etc. are already considered. When considered in conjunction with the data extraction process then, CAMELOT may actually reduce the amount of time a review team uses for critical appraisal. Furthermore, existing critical appraisal tools often ask users to answer yes/no or check a box for each critical appraisal criterion. This type of approach is insufficient in the context of a GRADE-CERQual assessment. Thus, the potential extra time used with CAMELOT for noting concerns is critical to adequately addressing the GRADE-CERQual methodological limitations domain.

CAMELOT is intentionally flexible to accommodate the types of primary studies being assessed and the type of QES being conducted. Review authors will need to follow the guidance and novices may require training to use it appropriately. CAMELOT encourages review authors to engage with included studies and become familiar with, and be thoughtful regarding, the choice of, and rationale for, the study design and methods, and data derived from the studies. By incorporating CAMELOT into the data extraction phase, review teams can consider elements of the design and conduct of the study and how they fit together earlier within the review process.

BOX 2. Next steps and research agenda for CAMELOT.

As with all tool development, CAMELOT will further evolve over time. The following activities are planned to further develop and refine the approach:

- Further refinement of guidance and examples based on further user testing by review authors from different disciplines, with different levels of experience to explore possible revisions or adaptations for different types of primary studies, different QES methodologies, and whether current guidance is sufficient for review teams with less experience.
- Additional user testing and piloting to further enhance integration of CAMELOT with GRADE-CERQual.
- Exploring possibilities for integration with different qualitative synthesis tools (e.g., iSoQ [44] data extraction and analysis software).

A deliberate intended impact of CAMELOT is the improvement of the design and conduct of primary qualitative studies. CAMELOT assessments are dependent on adequate reporting of the domains in primary studies, especially related to domains such as theory and equity, diversity, and inclusion that have not traditionally been prioritized in reports of qualitative studies. In CAMELOT, the definition of equity, diversity, and inclusion builds upon existing literature on equity, diversity, and inclusion [36–43].

Given the evolving nature of CAMELOT, we have established a website where guidance and examples will be updated as we gather experience from users in applying the approach (see camelotapproach.com). We encourage those interested in using CAMELOT to visit this page for the most up-to-date guidance.

5.1 | Strengths and limitations

The most obvious strengths of the CAMELOT approach are the use of evidence to inform domain development, as well as extensive and prolonged engagement with relevant methodological experts and stakeholders (including end users) to develop the domains, definitions, and overall CAMELOT approach. Furthermore, while CAMELOT was developed to support review authors and with GRADE-CERQual in mind, it can be used alone to assess the methodological limitations of a primary study that is not party of a QES. However, CAMELOT has not been widely piloted or validated and any guidance provided in this report is subject to revision and change (see Box 2 for next steps in the development of CAMELOT). Furthermore, fewer panelists participated in the Delphi than expected and because of this we may have missed ideas and opinions from potentially key informants. We hope to ameliorate this through extensive user testing.

5.2 | Patient and public involvement

Multiple and diverse stakeholders were involved in developing CAMELOT. We sought feedback from experienced methodological experts in Phase 1. Researchers and end users interested in qualitative research were invited to give feedback on the findings from the systematic review via workshops. The results of stage 3 (search for evidence) were posted on a publicly accessible webpage for comments from relevant researchers. Consensus involved multiple methodological experts and end users. Finally, user testing provided feedback from review authors with varying levels of experience of CAMELOT domains and guidance.

6 | CONCLUSION

CAMELOT fills a longstanding vacuum in the methodology of QES. It was developed and subject to initial testing using rigorous evidence-based methods. We encourage experienced review authors to start using CAMELOT in their Cochrane and non-Cochrane QESs when using GRADE CERQual and to provide feedback on their experience. CAMELOT also offers an appropriate tool of choice for QESs that do not include a GRADE-CERQual assessment. We also encourage novice review authors to work with more experienced colleagues to gain experience of using CAMELOT. The CAMELOT tool is likely to evolve as the CAMELOT working group collects feedback from review authors' experience of using the approach.

The CAMELOT Development Group

Chris Carmona, Chris Colvin, Danielle Kelly Pollock, Elham Shakibazadeh, Emma France, Heather Ames, Heid Nøkleby, Irene Hadeland, Irene Ljødal, James Parrott, Jo Thompson-Coon, Judith Carrier, Kate McMahon, Lars Jørun Langøien, Leif Stromwall, Martina Giltenane, Meghan Wainright, Özge Tuncalp, Pamela Rothpletz-Puglia, Rachel Kettle, Tarang Sharma, Tari Turner, Unni Gopinathan, and Ursula Griebler.

AUTHOR CONTRIBUTIONS

Heather M. Munthe-Kaas: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; project administration; visualization; writing—original draft; writing—review and editing. **Andrew Booth**: Conceptualization; data curation; formal analysis; funding acquisition; investigation; methodology; writing—review and editing. **Isolde Sommer**: Formal analysis; investigation; methodology; writing—review and editing. **Sara Cooper**: Formal analysis; investigation; methodology; writing—review and editing. **Ruth Garside**: Formal analysis; investigation; methodology; writing—review and editing. **Karin Hannes**: Investigation; writing—review and editing. **Jane Noyes**: Conceptualization; formal analysis; funding acquisition; investigation; methodology; writing—review and editing.

ACKNOWLEDGMENTS

The authors acknowledge the use of ChatGPT following initial drafting of the original manuscript for editorial support and take full responsibility for the content of this article. We would like to acknowledge the Cochrane Methods Innovation Fund (Project nr. 1328523) and the Norwegian Institute of Public Health for funding this work.

DATA AVAILABILITY STATEMENT

Any Supporting Information not included in the manuscript is available upon request.

ORCID

Heather M. Munthe-Kaas https://orcid.org/0000-0002-6051-9010
Andrew Booth https://orcid.org/0000-0003-4808-3880
Isolde Sommer https://orcid.org/0000-0003-3592-1507
Sara Cooper https://orcid.org/0000-0002-9627-0309
Ruth Garside https://orcid.org/0000-0003-1649-4773
Karin Hannes https://orcid.org/0000-0002-5011-3615
Jane Noyes https://orcid.org/0000-0003-4238-5984

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/cesm.12058.

REFERENCES

1. Langlois S, Goudreau J, Lalonde L. Scientific rigour and innovations in participatory action research investigating workplace learning in continuing interprofessional education. *J Interprof Care*. 2014;28(3):226-231.
2. Carroll C. Qualitative evidence synthesis to improve implementation of clinical guidelines. *BMJ*. 2017;356:1-4.
3. Glenton C, Lewin S, Lawrie TA, et al. Qualitative evidence synthesis (QES) for guidelines: paper 3–using qualitative evidence syntheses to develop implementation considerations and inform implementation processes. *Health Res Policy Syst*. 2019;17(1):74.
4. Lewin S, Booth A, Glenton C, et al. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implementation*. 2018;2018:1-10.
5. Wainwright M, Zahroh RI, Tunçalp Ö, et al. The use of GRADE-CERQual in qualitative evidence synthesis: an evaluation of fidelity and reporting. *Health Res Policy Syst*. 2023;21(1):77.
6. Lewin S, Glenton C, Munthe-Kaas H, et al. 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 Med*. 2015;12(10):e1001895.
7. Hannes K. Critical appraisal of qualitative research. In *Cochrane Qualitative and Implementation Methods Group Online Guidance*. Cochrane Qualitative and Implementation Methods Group; 2011.
8. Glenton C, Colvin CJ, Carlsen B, et al. Barriers and facilitators to the implementation of lay health worker programmes to improve access to maternal and child health: a qualitative evidence synthesis. *Cochrane Database Syst Rev*. 2013;2013(10):CD010414.
9. Karimi-Shahanjarini A, Shakibazadeh E, Rashidian A, et al. Barriers and facilitators to the implementation of doctor-nurse substitution strategies in primary care: a qualitative evidence synthesis. *Cochrane Database Syst Rev*. 2019;2019(4):CD010412.
10. Lins S, Hayder-Beichel D, Rücker G, et al. Efficacy and experiences of telephone counselling for informal carers of people with dementia. *Cochrane Database Syst Rev*. 2014;2014(9):CD009126.
11. Williams V, Boylan A-M, Nunan D. Critical appraisal of qualitative research: necessity, partialities and the issue of bias. *BMJ Evid Based Med*. 2020;25(1):9-11.
12. Carroll C, Booth A. Quality assessment of qualitative evidence for systematic review and synthesis: is it meaningful, and if so, how should it be performed? *Res Synth Methods*. 2015;6(2):149-154. doi:10.1002/jrsm.1128

13. Munthe-Kaas HM, Glenton C, Booth A, Noyes J, Lewin S. Systematic mapping of existing tools to appraise methodological strengths and limitations of qualitative research: first stage in the development of the CAMELOT tool. *BMC Med Res Methodol*. 2019;19(1):113.
14. Popay J. Moving beyond floccinaucinihilipilification: enhancing the utility of systematic reviews. *J Clin Epidemiol*. 2005;58(11):1079-1080. doi:10.1016/j.jclinepi.2005.08.004
15. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care*. 2007;19(6):349-357.
16. Walsh D, Downe S. Appraising the quality of qualitative research. *Midwifery*. 2006;22(2):108-119.
17. Burls A. What is critical appraisal? What is...? Series. Hayward Medical Communications. 2019:NPR09/1113.
18. Munthe-Kaas HM. *Exploring the Development of the GRADE-CERQual and TRANSFER Approaches to Enhance Accountability for Reasonableness in Evidence-Informed Decision-making Processes*. University of Oslo; 2020.
19. Moher D, Schulz KF, Simera I, Altman DG. Guidance for developers of health research reporting guidelines. *PLoS Med*. 2010;7(2):e1000217.
20. Booth A, Sommer I, Cooper S, et al. *The CAMELOT Plot*. Accessed September 21, 2023. http://thecamelotplot.pbworks.com/w/page/136970658/FrontPage2023
21. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. *Palliat Med*. 2017;31(8):684-706.
22. Higgins JPT, Altman DG, Gotzsche PC, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.
23. Turoff M, Linstone HA. The Delphi Method-Techniques and Applications. Addison Wesley Publishing Company. 2002.
24. Munthe-Kaas H, Sommer I, Noyes J, et al. *Development of the CAMELOT Approach for Considering Methodological Limitations of Qualitative Research in the Context of GRADE-CERQual and Qualitative Evidence Syntheses—Protocol (Version 1)*. Zenodo; 2023.
25. Witteman H, Dansokho S, Colquhoun H, et al. User-centred design and the development of patient decision aids: protocol for a systematic review. *Syst Rev*. 2015;4(11):1-8.
26. Rosenbaum S. *Improving the User Experience of Evidence: A Design Approach to Evidence-Informed Health Care*. The Oslo School of Architecture and Design; 2010.
27. Rosenbaum S, Oxman M, Oxman AD, et al. *Human-Centred Design Development of Informed Health Choices (IHC) Learning Resources for Secondary School Students: Protocol*. IHC; 2019.
28. Rosenbaum SE, Glenton C, Nylund HK, Oxman AD. User testing and stakeholder feedback contributed to the development of understandable and useful summary of findings tables for cochrane reviews. *J Clin Epidemiol*. 2010;63(6):607-619.
29. Ericsson KA, Simon HA. Verbal reports as data. *Psychol Rev*. 1980;87(3):215-251.
30. Kuniavsky M. *Observing the User Experience: A Practitioner's Guide to User Research*. Elsevier; 2003.
31. Rosenbaum S. *Improving the User Experience of Evidence: A Design Approach to Evidence-Informed Health Care*. Oslo College of Architecture and Design; 2010.
32. Rosenbaum S, Moberg J, Chesire F, et al. Teaching critical thinking about health information and choices in secondary schools: human-centred design of digital resources. *F1000Research*. 2023;12:481.
33. Munthe-Kaas HM, Booth A, Noyes J. Assessing methodological strengths and limitations. In: Noyes J, Harden A, Ames H, eds. *Cochrane-Campbell Handbook for Qualitative Evidence Synthesis*. Wiley; 2024.
34. Abras C, Maloney-Crichmar D, Preece J. *User-Centred Design. Bainbridge, W Encyclopedia of Human-Computer Interaction*. Sage Publications; 2004.
35. Linstone HA, Turoff M. *The Delphi Method*. Addison-Wesley Reading; 1975.
36. Oliver S, Kavanagh J, Caird J, et al. *Health Promotion, Inequalities and Young People's Health: A Systematic Review of Research*, 2008.
37. O'Neill J, Tabish H, Welch V, et al. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *J Clin Epidemiol*. 2014;67(1):56-64.
38. Welch V, Petticrew M, Tugwell P, et al. PRISMA-Equity 2012 extension: reporting guidelines for systematic reviews with a focus on health equity. *PLoS Med*. 2012;9(10):e1001333.
39. Welch VA, Norheim OF, Jull J, et al. CONSORT-Equity 2017 extension and elaboration for better reporting of health equity in randomised trials. *BMJ*. 2017;359:1-14.
40. Kelly C, Dansereau L, Sebring J, et al. Intersectionality, health equity, and EDI: what's the difference for health researchers? *Int J Equity Health*. 2022;21(1):182.
41. HughsonWoodward-Kron, J, Woodward-Kron R, Parker A, et al. A review of approaches to improve participation of culturally and linguistically diverse populations in clinical trials. *Trials*. 2016;17:263.
42. Woodland L, Blignault I, O'Callaghan C, Harris-Roxas B. A framework for preferred practices in conducting culturally competent health research in a multicultural society. *Health Res Policy Syst*. 2021;19:24.
43. Secules S, McCall C What research can DO: rethinking qualitative research designs to promote change towards equity and inclusion. *Stud Eng Educ*. 2023;4(1):26-45.
44. GRADE-CERQual Interactive Summary of Qualitative Findings (iSoQ). [program]. 1.0 version. Norwegian Institute of Public Health. (Developed by the Epistemonikos Foundation, Megan Wainwright Consulting and the Norwegian Institute of Public Health for the GRADE-CERQual Project Group). 2023.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Munthe-Kaas HM, Booth A, Sommer I, et al. Developing CAMELOT for assessing methodological limitations of qualitative research for inclusion in qualitative evidence syntheses. *Cochrane Ev Synth*. 2024;2:e12058. doi:10.1002/cesm.12058