


ORIGINAL RESEARCH

Synthesis of guidance available for assessing methodological quality and grading of evidence from qualitative research to inform clinical recommendations: a systematic literature review

Mandeep Sekhon ¹, Annette de Thurah ^{2,3}, George E Fragoulis ⁴, Jan Schoones,⁵ Tanja A Stamm ⁶, Theodora P M Vliet Vlieland ⁷, Bente Appel Esbensen ^{8,9}, Heidi Lempp,¹⁰ Lindsay Bearne,¹ Marios Kouloumas,¹¹ Polina Pchelnikova,¹² Thijs Willem Swinnen ¹³, Chris Blunt,¹⁴ Ricardo J O Ferreira ^{15,16}, Loreto Carmona ¹⁷, Elena Nikiphorou ¹⁸

To cite: Sekhon M, de Thurah A, Fragoulis GE, *et al.* Synthesis of guidance available for assessing methodological quality and grading of evidence from qualitative research to inform clinical recommendations: a systematic literature review. *RMD Open* 2024;**10**:e004032. doi:10.1136/rmdopen-2023-004032

Received 3 January 2024
Accepted 14 March 2024



© Author(s) (or their employer(s)) 2024. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

Correspondence to
Dr Mandeep Sekhon;
msekhon@sghul.ac.uk

ABSTRACT

Objective To understand (1) what guidance exists to assess the methodological quality of qualitative research; (2) what methods exist to grade levels of evidence from qualitative research to inform recommendations within European Alliance of Associations for Rheumatology (EULAR).

Methods A systematic literature review was performed in multiple databases including PubMed/Medline, EMBASE, Web of Science, COCHRANE and PsycINFO, from inception to 23 October 2020. Eligible studies included primary articles and guideline documents available in English, describing the: (1) development; (2) application of validated tools (eg, checklists); (3) guidance on assessing methodological quality of qualitative research and (4) guidance on grading levels of qualitative evidence. A narrative synthesis was conducted to identify key similarities between included studies.

Results Of 9073 records retrieved, 51 went through to full-manuscript review, with 15 selected for inclusion. Six articles described methodological tools to assess the quality of qualitative research. The tools evaluated research design, recruitment, ethical rigour, data collection and analysis. Seven articles described one approach, focusing on four key components to determine how much confidence to place in findings from systematic reviews of qualitative research. Two articles focused on grading levels of clinical recommendations based on qualitative evidence; one described a qualitative evidence hierarchy, and another a research pyramid.

Conclusion There is a lack of consensus on the use of tools, checklists and approaches suitable for appraising the methodological quality of qualitative research and the grading of qualitative evidence to inform clinical practice. This work is expected to facilitate the inclusion

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ The importance of qualitative research in informing clinical practice is widely recognised.
- ⇒ Various tools exist to assess the quality of studies using qualitative methodologies.
- ⇒ There is lack of clarity on how to best use qualitative evidence in formulating clinical recommendations within European Alliance of Associations for Rheumatology (EULAR).
- ⇒ This includes a lack of guidance on grading the quality of the included qualitative studies and assessment of the strength of recommendations based on qualitative studies.

of qualitative evidence in the process of developing recommendations at EULAR level.

INTRODUCTION

There has been increasing recognition over time on the role of qualitative research, including in complementing quantitative, epidemiological research and informing clinical practice.^{1–5} In recent years, the patients' and carers' perspectives, specifically on the acceptability and feasibility of interventions and services, have been considered in the development of clinical guidelines and recommendations.^{6–8} Additionally, it is becoming more widely acknowledged that qualitative research has a significant role to

WHAT THIS STUDY ADDS

- ⇒ The findings from this systematic literature review provide clarity on the various tools available to assess the quality of qualitative studies. In addition, no validated hierarchy of evidence exists that can integrate the evidence from qualitative research to inform the development of clinical recommendations.
- ⇒ Our findings indicate the need to advance the field of assessing evidence from qualitative research to develop explicit and comprehensive criteria for downgrading and upgrading quality of evidence ratings, to inform the development of clinical recommendation within EULAR.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ Qualitative research often reflects the direct patient voice, which should always be considered in clinical decision practice (shared decision-making) and research.
- ⇒ There is an urgent need for the further development of tools and approaches to assess and grade evidence from qualitative studies, so that such evidence can be appropriately incorporated into clinical practice guidelines and recommendations.
- ⇒ Can help standardise and improve the use of qualitative research for the development of recommendations that are in alignment with the EULAR Standardised Operating Procedures using the Oxford hierarchy.
- ⇒ Can encourage the use of qualitative research (raising its profile within the community), especially when it comes to including the patient perspective to recommended interventions and balance benefits and harm.

play in addressing issues and concerns with complex interventions, and the guideline development process.⁸

Depending on the nature of the topic and research question (RQ), a range of different qualitative methods may be applied. These include structured interviews (eg, following a topic guide closely, where all participants are asked exactly the same questions), semistructured interviews (eg, including a range of closed and open-ended questions, and the interviewer may adapt the topic guide according to the responses of the participants); focus groups (often held with 8–12 participants from the target population, where group dynamics and collective perceptions on a topic are required) and observations (where researchers observe participants to gain an insight into their experiences).^{9,10} Furthermore, many different methodologies are applied to analyse qualitative research (ie, thematic analysis, grounded theory, qualitative content analysis).^{9,10}

Yet, there is a recognised lack of a ‘gold standard’ quality appraisal tool, explicit frameworks or guidelines on how best to use qualitative evidence, to formulate clinical recommendations and guidelines.^{11,12} This includes a lack of guidance on grading the quality of qualitative studies and assessing the strength of recommendations based on qualitative studies.¹³ In addition, the literature remains scarce with most of the existing grading systems originating from quantitative research, and, consequently, no methods are available describing how best to

incorporate qualitative research into the evidence hierarchy.^{5,13}

Thus, there is an important unmet need since qualitative research often reflects the direct patients’/carers’ voices. Without a clear framework of how best to use qualitative research, there is a danger of suboptimal employment of all the available evidence when it comes to formulating clinical recommendations.

In the context of European Alliance of Associations for Rheumatology (EULAR), most recommendations are based on quantitative research and there is clear guidance on how to use the existing evidence to formulate these recommendations. However, within EULAR, the Oxford methodology is used for evidence synthesis,¹⁴ without explicit and practical guidance on how to use qualitative studies. It is important to highlight that EULAR produces a broad scope of guidelines and recommendations on specific medical interventions. These guidelines cover many different aspects of treatment as well as guidelines, which focus on addressing the needs of people with rheumatic and musculoskeletal diseases (eg, self-management, adherence to treatment, remote care and patient education). In many of these guidelines, there is a strong need for answering specific questions using qualitative research to understand the perceptions and experiences from the point of view of individuals with rheumatic and musculoskeletal diseases.

This recognised unmet need, prompted the undertaking of this work. Recognising the general lack of information on the inclusion of qualitative research in guideline development,¹⁵ this work aimed to provide practical guidance at EULAR level, that can help standardise the way qualitative research is used to inform the development of recommendations.

Specific objectives included:

1. to identify what guidance (tools, checklists, frameworks) exists to assess the methodological quality of qualitative research and
2. to understand if a grading system could be used in ranking the level of evidence based on qualitative research.

This work aimed to inform the latest EULAR Standardised Operating Procedures (SOP) for formulating recommendations.¹⁶

METHODS

Research questions

This systematic literature review (SLR) was conducted following the EULAR Standardised operating procedures¹⁴ and the protocol for this review was registered on www.researchregistry.com (ID number: reviewregistry1240). The scope of the literature was defined during a meeting convened as part of a EULAR project (HPR048) group to address how best to apply qualitative research when formulating clinical recommendations. During this meeting, two key RQs were identified, and these formed the basis for the SLR.

RQ1 focused on identifying what guidance exists (eg, tools, checklists, frameworks) to assess the methodological quality of qualitative research to inform clinical recommendations. *The domain being studied is the guidance available to assess the methodological quality in qualitative research studies, which have directly been applied to inform the development of clinical recommendations.*

RQ2 focused on understanding what methods exist to grade levels of evidence for qualitative research to inform clinical recommendations. *The domain being studied is the methods available to grade levels of evidence for qualitative research.*

It is important to note the purpose of our search strategy, and eligibility assessment was to identify and describe high-utility appraisal tools relevant to assessing the quality of qualitative research applied to inform the development of clinical recommendations (RQ1), rather than to provide a comprehensive description of all the tools in the literature for qualitative appraisal. Similarly, for RQ2, our focus was to identify methodologically how in the published literature qualitative research been 'graded' to inform the development of clinical recommendations. Ultimately, the question of interest was whether there is an acceptable hierarchy of evidence in qualitative research, as in the case of quantitative research.

Search methodology

The two RQs were broken down into two search strategies, which were defined in collaboration with a trained librarian (JS). For both search strategies, the following electronic databases were searched PubMed/Medline, EMBASE, Web of Science, COCHRANE, Emcare, PsycINFO, ERIC, Academic Search Premier, Sociological Abstracts, ProQuest Dissertations and Thesis Global for published and unpublished studies. Searches were completed from inception to 23 October 2020.

For search strategy 1, concepts searched in various permutations included 'level of evidence' and 'qualitative research' (online supplemental appendix). For search strategy 2, there were further three subsearch strategies, which focused on: (1) evidence synthesis/grading levels of evidence of qualitative studies; (2) evidence synthesis/grading of level of evidence and checklists/tools and (3) quality assessment and qualitative studies and checklist/tools (online supplemental appendix).

Study selection

Title, abstracts and full texts were assessed by two independent reviewers (MS and GEF) in the Rayyan software (<https://www.rayyan.ai/>). Eligibility criteria for both RQs were as follows: primary articles and guideline documents published in English, describing the (1) development; (2) application of validated tools (eg, checklists); (3) guidance on how to assess methodological quality of qualitative research and (4) guidance on how to grade levels of qualitative evidence. Opinion pieces and conference abstracts were excluded. Manual searches of the

reference lists of full-text articles were conducted. The reference lists of systematic reviews that were not explicitly related to our RQ were also searched for potentially relevant papers. Disagreements between the two reviewers were resolved through discussion with senior authors (EN, AdT).

Data extraction and risk of bias assessment

A standardised, prepiloted MS Excel data extraction was applied to extract data on the following: guidance/tool/checklist core characteristics (eg, authors, name of guidance document/checklist/tool, target audience, clinical context/rationale, evidence base); details on validation process and robustness (if available), and details on instructions provided on how to apply the tools to assess quality of evidence for qualitative research/grade levels of evidence of qualitative research.

For all included tools/checklists, we also reviewed whether research bodies or organisations endorsed or recommended if any of the included checklists/tools should be applied to assess the quality of qualitative research. The research bodies we considered included The Oxford Centre for Evidence Based Medicine¹⁷; National Institute for Health Research (NIHR)¹⁸; Cochrane¹⁹ and Scottish Intercollegiate Guidelines Network²⁰ and WHO.²¹ The strengths and weaknesses of the guidance documents available to assess quality of qualitative evidence and grade levels of qualitative research were noted and summarised.

Data analysis

A narrative synthesis was conducted manually to identify key themes within and between the included studies. Key themes focused on main aspects of the published guidance available to assess (a) methodological quality of qualitative research and (b) each of the methods described to grade levels of qualitative evidence. Areas which were considered to lack clarity across the guidance and methods available were also noted.

RESULTS

Study selection and study characteristics

Overall, across both searches, 9073 records were retrieved (figure 1). After deduplication and title/abstract screening, 51 full text articles were assessed for eligibility yielding 15 included articles. Reasons for exclusion were as follows: articles not focused on assessing quality of qualitative research (n=25); levels of evidence not applicable to qualitative research (n=7) and articles not relevant to developing clinical recommendations (n=4).

For RQ1, five articles were included, describing five tools (1) *The Society for Critical Care Medicine Family—Centred Care Guidelines*²²; (2) *Nursing Management of the Second Stage of Labour Evidence Based Clinical Practice Guidelines*²³; (3) *Jonna Briggs Institute Critical Appraisal of Qualitative Studies*²⁴; (4) *Critical Skill's Appraisal Programme (CASP)*²⁵ and (5) *the Modified CASP Checklist*²⁶ (table 1). All tools ranged from 10 to 30 items and evaluated research

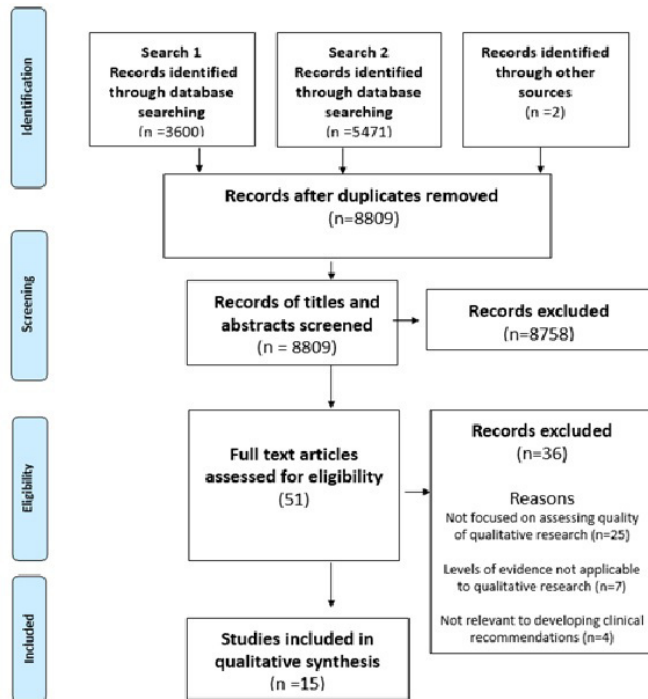


Figure 1 PRISMA diagram of included papers. PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

design, recruitment, ethical rigour, data collection and data analysis. Seven articles described one approach (Grading of Recommendations, Assessment, Development and Evaluation-Confidence in the Evidence from Reviews of Qualitative research (*GRADE-CERQual*)) on assessing confidence in synthesised findings within a qualitative systematic review. This approach is as known as *GRADE-CERQual*.^{27–33} This approach advised on the importance for assessing methodological limitations when considering overall confidence in qualitative synthesised findings (table 1).

For RQ2 (table 2), two articles were included, one described a qualitative hierarchy of evidence³⁴ and another described a research pyramid that included a section on qualitative research.³⁵

Summaries of included tools, approaches and hierarchy of evidence tools

Detailed data extraction for all the included papers are available in online supplemental file. Below we provide a brief description and summary of each of the included papers. Tables 1 and 2 provide key details of the main criteria.

RQ1: guidance (tools, checklists, frameworks) applied to assess the methodological quality of qualitative research

1. *The Society for Critical Care Medicine Family—Centred Care Guidelines*²²

Coombs and colleagues acknowledge that the use of qualitative research to inform professional guidelines within family centred care in intensive care units has been explored previously. These studies have focused

on exploring ‘which’ qualitative studies have been used in professional guidelines, rather than exploring ‘how and when’ qualitative research has been used in the guideline development process.²² Thus, the focus of their paper was to explore the importance, challenges and opportunities of using qualitative research to enhance development of clinical practice guidelines, specifically using recent guidelines published for family-centred care in the intensive care units as an example.

The family-centred care guidelines were developed based on their own experience of guidelines available in family centred care and a qualitative systematic review.²² The results of the systematic review were analysed via thematic analysis, and key themes produced, which then inform the Population, Intervention, Comparison, Outcome and key areas to include in the guidelines. Three experts in qualitative research and data analysis reviewed the tool prior to use and developed guidance for the guideline taskforce on how to apply the tool.

The tool includes 15 items (1) population; (2) age group; (3) methodology; (4) sampling method; (5) data collection method; (6) transferability/scale of population; (7) research design and methodology consistent with aim; (8) ethics review or IRB approval declared; (9) member checking/participant feedback; (10) coding analysis method description; (11) depth of reporting; (12) sample size methodology reported; (13) consistency; (14) paraphrased results; (15) major themes (aspects of FCC addressed). All items included various response options (see online supplemental file for full details of response options).

Each taskforce member reviewed between 8 and 10 papers. Once all reviews were complete, appraisal sheets were checked by the taskforce team leaders. Key findings from individual studies were subjected to constant comparison and thematic analysis. Common themes were identified and discussed among the guideline taskforce and validated by patients and families engaged in the study.

2. *Nursing Management of the Second Stage of Labour Evidence-Based Clinical Practice Guidelines*²³

Cesario *et al*²³ describe the process by which members of an Association of Women’s Health, Obstetric and Neonatal Nurses evidence-based clinical practice guideline development team created a scoring system based on specific criteria, for the evaluation of qualitative studies. Methods included an initial literature review on methods evaluating qualitative studies to inform the development of the criteria for the scoring tool were based on evaluating criteria for qualitative research proposed by Burns and Grove.³⁶

The scoring tool was designed to place qualitative studies included within their systematic review on specific levels of evidence (based on the overall quality rating). There are 30 criteria items phrased as questions, broken down into five main categories in the tool : (1)

Table 1 Guidance (tools, checklists, frameworks, approaches) applied to assess the methodological quality of qualitative research

	Society for Critical Care Medicine Family Centred Care Guidelines Coombs <i>et al</i> ²²	Nurse Management Labour Evidence-Based Clinical Practice Guidelines Cesario <i>et al</i> ²³	JBI Checklist Lockwood <i>et al</i> ²⁴	Critical Skills Appraisal Programme (CASP) CASP ²⁵	Modified CASP Long <i>et al</i> ²⁶	GRADE CERQual Lewin <i>et al</i> ²⁷
Country	New Zealand	USA	Australia	UK	UK	Norway
Study type	Case study	Literature review to inform development of criteria	Checklist	Checklist	Checklist modification of original CASP tool	Part of a series of papers for GRADE-CERQual approach
Guidance type	Tool	Tool	Checklist	Checklist	Checklist	Approach
Validated	Yes	Yes	Yes	Yes	Yes	Yes
Support for research bodies	No	No	NIHR ¹⁸	NIHR ¹⁸ (preferred checklist to apply) COCHRANE ¹⁹ (not in main handbook, but in supporting word document) WHO ²¹	COCHRANE ¹⁹ (not in main handbook, but a supporting word document) WHO ²¹	NIHR ¹⁸ COCHRANE ¹⁹ WHO ²¹
Clear criteria to determine overall quality of evidence	No	Yes Total score out of 30, corresponds to three quality levels	No	No	No	No
Similar items across all tools based around	1. Items on sampling methods, research design and methodological consistency, sample size, ethics review, analysis method, consistency- does the data match conclusions 2. N/A	1. Items within 'methodological congruence' include subcategories of rigour in procedural rigour, ethical rigour, and confirmability 2. Theoretical connectedness— <i>for example, are theoretical concepts adequately defined and/or validated?</i>	1. Important to assess for methodological limitations for each review finding. 2. N/A	1. Is the methodology appropriate— research design, recruitment strategy, data collection, data analysis 2. N/A	1. Is the methodology appropriate— research design, recruitment strategy, data collection, data analysis 2. New item: <i>Are the study's theoretical underpinnings clear consistent and conceptually coherent?</i>	1. Four components that should be considered when deciding on the overall confidence to place in individual review findings (1) <i>methodological limitations,</i> (2) <i>coherence,</i> (3) <i>adequacy of data,</i> (4) <i>relevance.</i> N/A
GRADE-CERQual, Grading of Recommendations, Assessment, Development and Evaluation-Confidence in the Evidence from Reviews of Qualitative research.						

Table 2 Methods applied to grade levels of evidence based on qualitative research

	A hierarchy of evidence for assessing qualitative research Daly et al ³⁴	Research pyramid: a new evidence-based practice model for occupational therapy Tomlin and Borgetto ³⁵
How to grade levels of qualitative search?	Via, a 'hierarchy of evidence for assessing qualitative research'	Via, a 'research pyramid'
Objective of guidance	Outline explicit criteria for assessing the contribution of qualitative empirical studies in health and medicine, leading to a hierarchy of evidence specific to qualitative methods.	Separating the evidence-level criteria of internal and external validity, incorporating explicitly the evidence provided by qualitative studies, and retaining the critical notion of rigour, developed a new pyramidal evidence model
Key assumptions of method	'Generalisability' If ideal generalisable study is realised, we should have a research study that provides evidence that is secure, evidence that a reader can trust, and evidence that a policy maker or practitioner can use with confidence as the basis for decision making and policy generation	'Rigour' Considered how different studies of the same methodology type have differing amounts of rigour and thus provide different amounts of confirmable evidence of intervention effectiveness and applicability. A hierarchy could be established on each side of the triangle, that is, a vertical dimension of rigour could be added
Figure	<p>Fig. 1. A hierarchy of evidence-for-practice in qualitative research—study types and levels.</p>	<p>Figure 1. Research Pyramid. Note: Meta = meta-analyses.</p>
Assessment criteria for qualitative studies/details on types of qualitative research	Defining a theoretical framework Specifying the sampling process Describing the methods for data collection and analysis Drawing research conclusions	Qualitative research (side) <ol style="list-style-type: none"> 1. Meta synthesis of related qualitative studies 2. Group qualitative studies with more rigour (a b c) 3. Group qualitative studies with less rigour (a, b, c) <ol style="list-style-type: none"> a. prolonged engagement with participants b. Triangulation of data (multiple sources) c. Confirmation of data analysis and interpretation (peer and member checking) 4. Qualitative studies with a single informant

descriptive vividness (eg, is essential descriptive information included?); (2) methodological congruence (eg, are all elements or steps of the study presented accurately and clearly?); (3) analytical preciseness (eg, are the study conclusions based on the data gathered?); (4) theoretical connectedness (eg, are the theoretical concepts adequately defined and/or validated by data?) and (5) heuristic relevance (eg, are the findings relevant to nursing practice?).

Each of the items within the five categories should be scored from 0 to 3, with a score of 0=no evidence that the criteria have been met ≤25% of criteria met, 1=poor, 25%–49% criteria met, 2=fair, 50%–74%

criteria met and 4=good, 75%–100% criteria met. The final quality of evidence rating is based on the total quality score across the 30 items, with three quality levels: **QI**=total score of 22.5–30 indicates that 75%–100% of the criteria had been met; **QII**=total score of 15–22.4 indicates that 50% to 74% of the criteria had been met and **QIII**=total score of less than 15 indicates less than 50% of the criteria were met. Two teams served as first evaluators of the validity of the tool and a subsequent guideline development team has used the tool too.

3. *Jonna Briggs Institute Critical Appraisal of Qualitative Studies*²⁴

The Joanna Briggs Institute (JBI) is an international organisation dedicated to the promotion and adoption of evidence-based practice, offering a selection of critical appraisal tools. The qualitative appraisal tool has been designed to quality appraise papers for consideration in a qualitative literature review. The tool includes 10 critical appraisal questions, with four response categories (yes, no, unclear, not applicable) and an overall appraisal decision if the paper rated should be included, excluded, seek further information and a space for comments (eg, including reasons for exclusion).

4. *Critical Skills Appraisal Programme CASP*²⁵

The Critical Appraisal Skills Programme (CASP) qualitative tool is the most used tool for quality appraisal in health-related qualitative evidence syntheses, with endorsement from the Cochrane Qualitative and Implementation Methods Group.³³ The CASP tool is broken down into three sections, with 10 items that focus on three broad issues authors propose need to be considered when evaluating a qualitative study. These include section (A) are the results of the study valid?; section (B) What are the results and section (C) Will the results help locally. Response options for each of the items include yes, no and can't tell, with a space for comments to provide reasons for your answers. Several italicised prompts are given after each question, which are designed to remind researchers why the question is important. The first two questions are screening questions ((1) Was there clear statement of the aims of the research and (2) Is the qualitative methodology appropriate?). Researchers are advised that if the answer to both questions is 'yes', it is worth proceeding with the remaining questions.

5. *Modified CASP Checklist*²⁶

The modified CASP tool was developed by Long *et al.*,²⁶ to consider issues related to the suitability and usability of the CASP tool for quality appraisal in qualitative evidence synthesis. The authors reflect on their practical experience of using the original CASP tool in systematic reviews and qualitative evidence synthesis. The main modification proposed by authors includes an item that focuses on the study's underlying theoretical, ontological and epistemological framework. This item also includes a number of hints to consider when answering the item (see online supplemental file). Another modification includes a fourth response option to the previous response options included in the original CASP tool, 'somewhat'. The authors describe the 'somewhat' option to mean 'to some extent' or 'partly' for use when we deemed that the primary study authors had reported a reasonable attempt at fulfilling a particular quality domain but had clear strengths and limitations.²⁶ In terms of validation, the authors have applied the modified CASP tool to 10 papers included in a systematic review (see Long *et al.*²⁵ for details).

6. *GRADE CerQUAL*²⁷⁻³³

There were seven papers included in our review, which describe one approach on assessing confidence in synthesised findings within a qualitative systematic review. This approach is known as GRADE-CERQual.²⁷⁻³³ This approach was developed in line with the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) for quantitative research. The GRADE-CERQual approach provides guidance for assessing how much confidence to place in findings from systematic reviews of qualitative research (or qualitative evidence syntheses). The approach has been developed to support the use of findings from qualitative evidence syntheses in decision-making, including guideline development and policy formulation. The CERQual approach aims to 'systematise the process of assessing confidence in the evidence from qualitative evidence syntheses and make these assessments explicit and transparent'.²⁷

According to this approach, there are four components that need to be considered when summarising the overall quality of the body of evidence based on qualitative data: (1) *methodological limitations*, (2) *coherence*, (3) *adequacy of data* and (4) *relevance*.

The assessments of the four components collectively contribute to an overall assessment of whether findings from a qualitative evidence synthesis provide a reasonable representation of the health or social care issues, intervention or programme (phenomenon) of interest.

RQ2: methods applied to grade levels of evidence based on qualitative research

Table 2 provides details of the two articles that describe methods applied to grade levels of evidence based on qualitative research. One article described a qualitative hierarchy of evidence³⁴ and another described a research pyramid that included a section on qualitative research.³⁵

1. *A hierarchy of evidence-for-practice in qualitative research*³⁴

Daly and colleagues outline explicit criteria for assessing the contribution of qualitative empirical studies in health and medicine, leading to their hierarchy of evidence specific to qualitative methods.³⁴ The emphasis of the hierarchy is to rank qualitative study designs to assess the quality of evidence that a well-conducted study offers when practical decisions to inform clinical practice or policy need to be made. Criteria include (1) defining a theoretical framework for the study; (2) specifying a sampling process; (3) describing the methods of data collection and analysis and (4) drawing research conclusions.

According to this hierarchy, single case studies are considered of lowest evidence, followed by descriptive studies that may provide useful quotations but often lack detailed analysis. Conceptual studies tend to take greater significance, incorporating analysis based on conceptual themes, but lack of diversity in the sample can be a disadvantage. Generalisable studies that use conceptual frameworks that incorporate a diversified

sample and consider the use of all appropriate data provide the best evidence-for-practice. According to the authors, ‘generalisable studies can lead to provision of evidence that is secure, trustable and of use to policymakers or practitioners’.³⁴

2. *Research pyramid: a new evidence-based practice model for occupational therapy*³⁵

The research pyramid was developed as a revised model of evidence-based medicine to be applied with models of clinical reasoning within the field of occupational therapy.³⁵ The pyramid consists of a triangle with different study designs (experimental research, outcome research and qualitative research) placed in order/tiers of internal and external validity.³⁵ The qualitative side of the pyramid is placed in order of study designs that focus on detailed analysis and triangulation of data from different sources. For instance, a meta-synthesis of relate qualitative studies on a similar topic is placed at the top of the pyramid, and qualitative studies that use a single informant are placed at the bottom of the pyramid.

DISCUSSION

This systematic literature review highlights that various tools exist to assess the quality of qualitative studies. However, no validated hierarchy of evidence exists to integrate the evidence from qualitative research to inform the development of clinical recommendations. This is relevant in the context of formulating recommendations, which is how this work began in the first place: a systematic review of available evidence, to inform the EULAR SOP on the development of recommendations, specifically when dealing with qualitative studies.¹⁶ As in the case for quantitative research, a similar ‘standard’ was necessary for qualitative research, to ensure appropriate handling of qualitative evidence when formulating recommendations.

We found five tools (including checklists) that had been applied to assess the quality of qualitative studies. All tools had similarities in evaluating the research design of the qualitative study, recruitment, ethical rigour, data collection and data analysis.^{22–26}

One of the tools used for appraising the quality of qualitative studies had a ‘total score’ to determine the overall level of quality.²³ However, it can be argued that this approach could be problematic as it may overlook specific limitations of the study. This issue has also been found in quantitative tools, resulting in Cochrane revising their Risk of Bias tool for randomised controlled trials into a domain-based instrument. The remaining four were based on evaluation of domains in which critical assessments are made separately for different domains.^{22–24–26} For the purpose of EULAR, a more domain-based approach to appraising qualitative studies may be more appropriate.

The GRADE-CERQual approach was also identified in our review, where researchers aim to determine the

level of confidence to place in the findings from qualitative evidence synthesis/systematic reviews of qualitative research.^{27–33} In this approach, confidence is determined by considering four key components: (1) methodological limitations; (2) coherence; (3) adequacy of the data and (4) relevance.^{29–31} In line with the GRADE process for quantitative research, it is important to note the aim of the GRADE CERQual approach is to provide transparency in the process of moving from evidence to recommendation by evaluating these four key components. Users are advised to make a written note explaining their decision, and then reread and evaluate the decisions for all four components to reach an overall confidence assessment. As a ‘process’, it emphasises the usefulness of qualitative evidence in decision-making but still highlights the need for qualitative evidence not to be used alone, when making decisions about interventions.

It is worth noting that despite some similarities in the criteria for each of the tools, it is still difficult to compare the validity. Thus, it is important that any tool that is recommended for evaluating qualitative evidence is clear, practical and easy to use. This task force, with members representing various key stakeholders including patients, clinicians and methodologists, aimed at providing simple guidance on how to use qualitative research for the purpose of standardising the approach to the development of EULAR recommendations, when formulated questions can primarily be answered by qualitative research. It was beyond the scope of this project to develop a new guideline.

With regards to methods that can be applied to summarise the quality of qualitative research for clinical recommendations, our review found one qualitative evidence hierarchy³⁴ and a research pyramid that included a section on qualitative research.³⁵ Both approaches share some similarities in their criteria for assessing qualitative research studies. Daly *et al*³⁴ indicate at the top of the hierarchy are ‘generalisable studies’, which include evidence from qualitative systematic reviews. Similarly, within Tomlin and Borgetto’s³⁵ research pyramid, the qualitative meta synthesis is placed at the top of the pyramid. Thus, while both approaches may not be mutually exclusive, both seem to suggest that qualitative studies with the most ‘evidence’ should be considered as the top level, similar to meta-analysis of randomised controlled trials in the quantitative evidence hierarchy of evidence.

The lack of a hierarchy of evidence to grade qualitative research has not come by surprise, simply because such hierarchies do not seem to provide any additional value or guidance as to the quality of the evidence available. This is particularly so since qualitative research addresses aspects such as people’s experiences and factors that may influence barriers and facilitators to implementation, or acceptability and feasibility aspects.

As part of our data analysis, we also considered which of the included tools and approaches were endorsed or recommended by research bodies and funding

organisations. The NIHR¹⁸ recommended the use of two of the checklists as their preferred method of assessment, JBI²⁴ and CASP²⁵ and indicated when completing a qualitative evidence synthesis, the GRADE-CERQual approach can be used.²⁷ Similarly, Cochrane recommends the use of GRADE-CERQual approach in their handbook on qualitative evidence synthesis,¹⁹ and in a supporting document (not the actual handbook)³⁷ mention the use of the original CASP checklist²⁵ and the modified CASP checklist²⁶ in evaluating the quality of qualitative studies. The WHO also provides guidance on incorporating qualitative research when developing clinical guidelines.²¹

Strengths and limitations

Our review has strengths and limitations. To our knowledge, our SLR specifically focused on the quality of qualitative appraisal tools and is an updated review from previous similar reviews, which have centred around how qualitative evidence synthesis and appraisal tools have been applied in the development of clinical guideline.^{8,13,15}

Our review has considered the availability of guidance (in the form of tools, checklists, and frameworks) to assess the methodological quality of qualitative research and the methods applied to grade levels of evidence from qualitative research to inform the formulation of clinical recommendations and guidelines. Strengths include our use of systematic methods, screening and data extraction was completed by two review authors.

Limitations of our study include that our search strategy was completed in 2020, and that we did not incorporate guidelines for reviewers about reporting qualitative research, which was the focus of a similar published review.³⁸ This is because the intention of this review was to inform the EULAR SOP about developing clinical practice guidelines¹⁶ and, therefore, this was not considered relevant for this review.

Implications

This review highlights the complexity of the topic and helps inform on specific tools that can be used in the context of EULAR recommendations, to standardise the process. Furthermore, it is important to notice that EULAR guidelines cover many different aspects of treatment and care of people with rheumatic and musculoskeletal diseases—including, for example, self-management, adherence to treatment, remote care and patient education. In many of these guidelines, there is a strong need for answering specific questions using qualitative research.

CONCLUSION

Various checklists exist to assess the methodological quality of qualitative research. However, no hierarchy of evidence is applicable to categorise the level of evidence, in contrast with quantitative research. The findings of this review suggest that the quality of papers retrieved from systematic literature reviews needs to be consistently and rigorously checked, using an established checklist that

assesses the methodological appropriateness, research design, recruitment strategy, data collection, data analysis and theoretical underpinnings. By presenting these tools and checklists, we hope to encourage EULAR task forces to include qualitative research into the body of evidence of the developed guidelines where this methodology will in fact most precisely provide answers to the question.

A regular review of the evidence in this rapidly evolving field, with input from qualitative research methodology experts and leaders in the field, will be necessary to remain up to date and to give qualitative research the deserved attention in the development of relevant guidelines.

Author affiliations

¹Population Health Research Institute, St George's University of London, London, UK

²Department of Rheumatology, Aarhus University Hospital, Aarhus, Denmark

³Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

⁴Rheumatology, Army Shared Fund Hospital, Athens, Greece

⁵Leiden University Medical Center (LUMC), Directorate of research Policy, Leiden, Netherlands

⁶Section for Outcomes Research, Medical University of Vienna, Vienna, Austria

⁷Dept of Orthopaedics, Rehabilitation and Physical Therapy, Leiden University Medical Center (LUMC), Leiden, Netherlands

⁸Copenhagen Center for Arthritis Research (Copecare), Rigshospitalet, Glostrup, Denmark

⁹Department of Clinical Medicine, University of Copenhagen, Copenhagen, Denmark

¹⁰Centre for Rheumatic Diseases, Department of Inflammation Biology, School of Immunology and Microbial Sciences, Faculty of Life Sciences & Medicine, King's College London, London, UK

¹¹Cyprus League Against Rheumatism, Aglantzia, Cyprus

¹²PARE Research Partner, EULAR, Zurich, Switzerland

¹³Division of Rheumatology, KU Leuven University Hospitals Leuven, Leuven, Belgium

¹⁴London School of Economics and Political Science, London, UK

¹⁵Nursing Research, Innovation and Development Centre of Lisbon (CIDNUR), Nursing School of Lisbon (ESEL), Lisbon, Portugal

¹⁶Instituto de Saúde Ambiental (ISAMB), Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal

¹⁷Instituto de Salud Musculoesquelética, Madrid, Spain

¹⁸Centre for Rheumatic Diseases, School of Immunology and Microbial Sciences, King's College London, London, UK

X Mandeep Sekhon @mandy_sekhon, Marios Kouloumas @kouloumasm, Ricardo J O Ferreira @FerreiraRJO and Loreto Carmona @carmona_loreto

Acknowledgements This manuscript reports in detail work previously presented in a poster format at the EULAR 2022 Congress, Copenhagen 1–4 June. The conference abstract has been published: MS, AdT, GEF, TAS, TPMV, BAE, HL, LB, MK, PP, TWS. POS1552-HPR—A Synthesis of Guidance Available for Assessing Methodological Quality and Grading of Evidence from qualitative research to Inform Clinical Recommendations: A Systematic Review.

Contributors Authors ADT and EN are co-grauntors for the study. All authors contributed intellectually to this systematic review, the writing, revisions and final manuscript.

Funding European Alliance of Associations for Rheumatology.

Competing interests EN has received speaker honoraria/participated in advisory boards for Celltrion, Pfizer, Sanofi, Gilead, Galapagos, AbbVie, Fresenius-Kabi, Lilly and holds research grants from Pfizer and Lilly.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Data are available upon reasonable request. All data relevant to this study are published in the article or in the supplementary files.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been

peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Mandeep Sekhon <http://orcid.org/0000-0002-5109-9536>
 Annette de Thurah <http://orcid.org/0000-0003-0103-4328>
 George E Fragoulis <http://orcid.org/0000-0003-4932-7023>
 Tanja A Stamm <http://orcid.org/0000-0003-3073-7284>
 Theodora P M Vliet Vlieland <http://orcid.org/0000-0001-6322-3859>
 Bente Appel Esbensen <http://orcid.org/0000-0001-5331-8221>
 Thijs Willem Swinnen <http://orcid.org/0000-0002-5289-1023>
 Ricardo J O Ferreira <http://orcid.org/0000-0002-2517-0247>
 Loreto Carmona <http://orcid.org/0000-0002-4401-2551>
 Elena Nikiphorou <http://orcid.org/0000-0001-6847-3726>

REFERENCES

- Flemming K, Booth A, Garside R, *et al*. Qualitative evidence synthesis for complex interventions and guideline development: clarification of the purpose, designs and relevant methods. *BMJ Glob Health* 2019;4:e000882.
- Noyes J, Booth A, Cargo M, *et al*. Cochrane qualitative and implementation methods group guidance series—paper 1: introduction. *J Clin Epidemiol* 2018;97:35–8.
- Curry LA, Nembhard IM, Bradley EH. Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation* 2009;119:1442–52.
- Tan TPY, Stokes T, Shaw EJ. Use of qualitative research as evidence in the clinical guideline program of the National Institute for health and clinical excellence. *Int J Evid Based Healthc* 2009;7:169–72.
- Jones R. Strength of evidence in qualitative research. *J Clin Epidemiol* 2007;60:321–3.
- Cooper K, Kirkpatrick P, Florida-James S. Incorporating qualitative evidence in clinical practice guidelines: a scottish perspective. *Int J Evid Based Healthc* 2019;17 Suppl 1:S6–8.
- Lewin S, Glenton C. Are we entering a new era for qualitative research? Using qualitative evidence to support guidance and guideline development by the World Health Organization. *Int J Equity Health* 2018;17:126.
- Noyes J, Booth A, Moore G, *et al*. Synthesising quantitative and qualitative evidence to inform guidelines on complex interventions: clarifying the purposes, designs and outlining some methods. *BMJ Glob Health* 2019;4:e000893.
- Moser A, Korstjens I. Series: practical guidance to qualitative research. Part 1: introduction. *Eur J Gen Pract* 2017;23:271–3.
- Cleland JA. The qualitative orientation in medical education research. *Korean J Med Educ* 2017;29:61–71.
- Melnik BM, Gallagher-Ford L, Fineout-Overholt E. *Implementing the evidence-based practice (EBP) competencies in healthcare: a practical guide for improving quality, safety, and outcomes*. Sigma Theta Tau, 2016.
- Samaan Z, Mbuagbaw L, Kosa D, *et al*. A systematic scoping review of adherence to reporting guidelines in health care literature. *J Multidiscip Healthc* 2013;6:169–88.
- Majid U, Vanstone M. Appraising qualitative research for evidence syntheses: a compendium of quality appraisal tools. *Qual Health Res* 2018;28:2115–31.
- van der Heijde D, Aletaha D, Carmona L, *et al*. Update of the EULAR standardised operating procedures for EULAR-endorsed recommendations. *Ann Rheum Dis* 2015;74:8–13.
- Wang YY, Liang DD, Lu C, *et al*. An exploration of how developers use qualitative evidence: content analysis and critical appraisal of guidelines. *BMC Med Res Methodol* 2020;20:160.
- EULAR Sops standard operating procedures for task forces. n.d. Available: https://www.eular.org/myUploadData/files/update_eular_sops_06_10_2022_web_vs.pdf
- The Oxford centre for evidence based medicine (CEMB), Available: <https://www.cebm.ox.ac.uk/>
- National Institute for Health Research (NIHR). Available: <https://www.nihr.ac.uk/>
- Cochrane. Available: <https://www.cochrane.org/>
- Scottish intercollegiate guidelines network (SIGN). n.d. Available: <https://www.sign.ac.uk/>
- Downe S, Finlayson KW, Lawrie TA, *et al*. Qualitative evidence synthesis (QES) for guidelines: paper 1—using qualitative evidence synthesis to inform guideline scope and develop qualitative findings statements. *Health Res Policy Syst* 2019;17:76.
- Coombs MA, Davidson JE, Nunnally ME, *et al*. Using qualitative research to inform development of professional guidelines: a case study of the society of critical care medicine family-centered care guidelines. *Crit Care Med* 2017;45:1352–8.
- Cesario S, Morin K, Santa-Donato A. Evaluating the level of evidence of qualitative research. *JOGNN* 2002;31:708–14.
- Jonna Briggs Institute Critical Appraisal of Qualitative Studies. Available: https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Qualitative_Research2017_0.pdf
- Critical Appraisal Skills Programme. CASP (qualitative) checklist. 2018. Available: <https://casp-uk.net/wp-content/uploads/2018/01/CASP-Qualitative-Checklist-2018.pdf>
- Long HA, French DP, Brooks JM. Optimising the value of the critical appraisal skills programme (CASP) tool for quality appraisal in qualitative evidence synthesis. *Research Meth Med Health Sci* 2020;1:31–42.
- Lewin S, Booth A, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings: introduction to the series. *Implement Sci* 2018;13:2.
- Lewin S, Bohren M, Rashidian A, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 2: how to make an overall cerqual assessment of confidence and create a summary of qualitative findings table. *Implementation Sci* 2018;13:11–23.
- Munthe-Kaas H, Bohren MA, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 3: how to assess methodological limitations. *Implement Sci* 2018;13:9.
- Colvin CJ, Garside R, Wainwright M, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 4: how to assess coherence. *Implement Sci* 2018;13:13.
- Glenton C, Carlsen B, Lewin S, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 5: how to assess adequacy of data. *Implementation Sci* 2018;13:43–50.
- Noyes J, Booth A, Lewin S, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 6: how to assess relevance of the data. *Implementation Sci* 2018;13:51–61.
- Booth A, Lewin S, Glenton C, *et al*. Applying GRADE-CERQual to qualitative evidence synthesis findings—paper 7: understanding the potential impacts of dissemination bias. *Implementation Sci* 2018;13:63–70.
- Daly J, Willis K, Small R, *et al*. A hierarchy of evidence for assessing qualitative health research. *J Clin Epidemiol* 2007;60:43–9.
- Tomlin G, Borgetto B. Research pyramid: a new evidence-based practice model for occupational therapy. *Am J Occup Ther* 2011;65:189–96.
- Burns N, Grove SK. *Understanding nursing research* (2nd Ed). Philadelphia: W. B. Saunders, 1999.
- The Cochrane qualitative and implementation methods group. n.d. Available: <https://methods.cochrane.org/qi/>
- Buccheri RK, Sharifi C. Critical appraisal tools and reporting guidelines for evidence-based practice. *Worldviews Evid Based Nurs* 2017;14:463–72.