

Measurement properties of appraisal tools for mixed methods research: A systematic review protocol

Sachiko Makabe ^{1,*}, Tomomi Suda ¹, Yuko Akagawa ¹, Michiko Abe ² and Hisako Kakai ²

¹ Department of Nursing, Akita University Graduate School of Health Sciences, Akita, Japan.

² School of International Politics, Economics, and Communication, Aoyama Gakuin University, Tokyo, Japan.

Abstract

Purpose: Evaluating the study quality for mixed methods research is more challenging than assessing that for monomethod research because of the former's greater complexity and procedural multiplicity—particularly, the integration of quantitative and qualitative components. Consensus on the appraisal tools of mixed methods studies is important to ensure that such studies are properly planned, conducted, reported, and reviewed. This review aims to provide a comprehensive overview of the measurement properties of appraisal tools developed to assess research quality in mixed methods research.

Method: Inclusion criteria are as follows: (1) The population under investigation must be researchers. (2) Instruments must provide items, such as design, to evaluate the quality of mixed methods studies. (3) Outcomes must focus on validity (content validity, structural validity) and reliability (internal consistency). (4) Finally, the study type must be measurement property research, and hence, tool development and evaluation studies are eligible. We will search for studies published and dissertations in English and Japanese from a variety of sources, including Scopus, MEDLINE, CINAHL, PsycINFO, and dissertation ERIC, ProQuest, CiNii (Japanese), Ichushi-Web (Japanese), and National Diet Library (Japanese). The duration of the search is not limited. All the citations will be screened and selected by two independent reviewers. Data extraction, quality assessment, and syntheses of included studies will be performed according to the Consensus-based Standards for the Selection of Health Measurement Instruments criteria.

Systematic review registration number: PROSPERO registration number: CRD42021238210.

Keywords: Assessment; Checklist; Critique; Design; Integration

1. Introduction

Unraveling the mechanisms of complex phenomena is often required in the health sciences [1]. Consequently, the use of mixed methods research has increased due to the limitation of the mono method. Mixed methods research [2] is not confined to the integration of qualitative and quantitative data collection and analysis. It also entails integration at philosophical as well as paradigmatic levels [3]. Further, mixed methods research involves data integration at multiple levels before drawing conclusive interpretations [4]. Combining the strengths of both quantitative and qualitative methods, it achieves to tackle complex questions more than using either approach alone. Therefore, mixed methods research is a valuable approach for researchers to address the multifaceted nature of the phenomena under investigation in health sciences by examining prevalence, factors, relationships, and effectiveness, and simultaneously capturing individual views, opinions, perceptions, and experiences.

* Corresponding author: Sachiko Makabe

Department of Nursing, Akita University Graduate School of Health Sciences, Akita, Japan.

Despite many advantages of mixed methods research, assessing the quality of studies using this approach is more challenging than examining that for monomethod research because of the former's greater complexity and procedural multiplicity—specifically, the integration of quantitative and qualitative components [5]. Consensus on the appraisal tools of mixed methods studies is important to ensure that such studies are properly planned, conducted, reported, and reviewed. There are nine established sets of appraisal tools for mixed methods studies [6-8], including an appraisal tool,[9] an evaluation rubric [10], and seven criteria checklists [2, 8, 11-15]. Understanding the available appraisal tools and their characteristics could be beneficial for various types of researchers, such as inexperienced mixed methods researchers, peer reviewers, and reviewers conducting systematic reviews. Identifying the most reliable and validated tool is the cornerstone of mixed methods research. Appraisal tools are used by inexperienced mixed methods researchers to maintain the quality of their research protocol, by peer reviewers to assess the quality of papers, and by reviewers conducting systematic reviews to help improve the efficiency and reliability of their work. Consequently, to evaluate the quality of mixed methods appropriately, a systematic review is needed to suggest recommendations for the usage and further development of mixed methods appraisal tools.

We conducted a preliminary search for existing reviews on mixed methods appraisal tools using the following databases: Cochrane Library, JBI Evidence Synthesis, Scopus, COSMIN Database of Systematic Reviews, and PROSPERO. We found that there have been no systematic reviews of mixed methods appraisal tools. Instead, several reviews have focused on the quality of mixed methods research. Fábregues et al. (2017) [13] reviewed, reported, and recommended ways to address quality in mixed methods research and derived three major findings: (1) The number of reviews on mixed methods research quality is increasing. (2) Researchers have used a core set of quality criteria to appraise mixed methods research. (3) Further investigation focusing on quality, consistency in terminology, and the identification of core quality criteria is needed. Younas et al. (2019) [15] reviewed and reported mixed methods research in nursing. They concluded that mixed methods research is still nascent in nursing literature and that researchers are still learning to address the challenges during its execution, analysis, and reporting. Further, Fábregues et al. (2020) [16] reviewed and reported the use of mixed methods in methodological research on palliative and end-of-life care. They concluded that the studies' overall reporting quality needed improvement because the authors generally failed to adequately describe their research design and provided minimal detail regarding the integration of quantitative and qualitative approaches. They provided recommendations to improve overall reporting quality but did not specifically discuss tools for assessing the risk of bias.

Furthermore, we conducted an initial limited search of Scopus to identify studies on our research topic (see Appendix I). As data examples for this review, validity and reliability studies of the latest appraisal tools are described as follows. The mixed methods appraisal tool (MMAT) was developed in 2011 following an extensive literature review, and its reliability and efficiency were tested [17]. Furthermore, the MMAT's usefulness was reported [18]. The MMAT's content validity was also reported [19]. Studies that meet the inclusion criteria for this systematic review exist. An exhaustive search of published and unpublished sources will be performed to locate more appraisal tools with measurement properties reports.

Information regarding the measurement properties of mixed methods appraisal tools is crucial to aid researchers in selecting the optimal tool for their specific purpose. This review expands on previous reviews by focusing specifically on validated instruments and documenting their measurement properties in detail. To this end, we follow the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines [20] for systematic reviews of measurement properties, which are recommended for reviews. Although these guidelines were originally used for patient-reported measurements, they can also be applied to other types of measurements. This review provides comprehensive and systematic information to guide the further development of well-validated mixed methods appraisal tools for use in research settings.

1.1 Review questions

This review aims to identify the measurement properties of mixed methods appraisal tools and attempts to answer the following review questions:

- What are the measurement properties of mixed methods appraisal tools for researchers assessing the quality of a mixed methods study (primary outcome)?
- What is the relevance of these tools for researchers in terms of their purpose, development process, application method, administration burden, number of items, and topics (secondary outcome)?

2. Inclusion criteria

2.1 Population

The population must only include researchers who need to assess the quality of mixed methods research—whether when reviewing or writing papers, or when planning or conducting research. This population can thus comprise all those who need to evaluate research quality, including master's-level, PhD-level, and post-doctoral students; full-time researchers; and professors. For settings, especially when conducting a systematic review, an assessment tool is required to evaluate the bias inherent in each selected study.

2.2 Instruments and construct

In this review, instruments are defined as providing criteria to evaluate the quality of mixed methods studies. An instrument is often referred to as an appraisal tool and can also be called a checklist, a critique form, an assessment tool, or an appraisal list. In the appraisal of research quality, the quality of mixed methods research, integration, and quantitative and qualitative components are evaluated. Additionally, terms such as “multi-methods [21]” and “multilevel mixed methods research [22]” are existed. Multi-methods incorporate more than one qualitative method and do not include any quantitative methods [21]. Multilevel mixed methods research is an advanced design of mixed methods research [22]. This review focuses on the basic design rather than advanced design. Basic core conditions are most important before classifying further advanced design condition. Therefore, in this review, the term “multi-methods” and “multilevel mixed methods research” are excluded.

2.3 Outcomes

This review considers studies that include the following outcomes for measurement properties:

2.3.1 *Validity (content validity, structural validity)*

Content validity is the most important measurement property for confirming that the tools are relevant [23]. A well-designed qualitative approach for item construction ensures content validity. For good content validity, the inclusion criteria are as follows:

- Relevant items in the construct interest,
- Relevant items in the target methodology of interest,
- Relevant items in the context of use of interest,
- Appropriate response options,
- An appropriate recall period,
- No missing key concepts,
- A tool instruction understood by the population of interest as intended,
- Items and response options understood by the population of interest as intended,
- Items appropriately worded, and
- The response option matching question.

Structural validity [24] refers to the extent to which the scores of an appraisal tool adequately reflect the dimensionality (domain) of assessing mixed methods quality.

2.3.2 *Reliability (internal consistency)*

The consistency [24] of scores upon measurement is vital for researchers who judge the quality of mixed methods papers equally—for example, different researchers assessing the same paper. Reliability is achieved by using kappa or intra-class correlation coefficients between each researcher's assessment results.

As not every study provides data for all measurement property outcomes, studies that report at least one outcome regarding validity or reliability are considered. Cross-cultural validity, hypotheses testing for construct validity, and responsiveness are not included because mixed methods appraisal tools are used by researchers to assess papers.

2.4 Type of study

This review considers studies that involve the development of appraisal tools or checklists to assess the quality of mixed methods research. Measurement property studies, such as tool development and evaluation studies, are eligible.

3. Methods

The proposed systematic review will be conducted in accordance with the COSMIN and Joanna Briggs Institute (JBI) systematic review guidelines [20, 25]. The review title has been registered at PROSPERO (registration number: CRD42021238210).

3.1 Search strategy

The search strategy aims to locate published studies and dissertations. An initial limited search of Scopus was conducted to identify studies on the topic (see Appendix I). The key search terms and combination forms are (“mixed methods”) AND (“appraisal tool” OR checklist OR “risk of bias assessment” OR “criteria”) AND (“reliability” OR “validity”). We will search for studies published in English and Japanese from a variety of sources because our research team is bilingual in English and Japanese. However, the duration of the search is not limited. The text words contained in the titles and abstracts of relevant studies and the index terms used to describe the studies are used to develop a full search strategy for each information source. The search strategy, including all the identified keywords and index terms, will be adapted for each database or information source. The reference lists of all included sources of evidence will be screened for additional studies.

3.2 Information sources

The following databases will be searched: Scopus, MEDLINE, CINAHL, PsycINFO, ERIC, ProQuest, CiNii (Japanese), Ichushi Web (Japanese), and National Diet Library (Japanese). ProQuest and National Diet Library index dissertations.

3.3 Study selection

Following the search, all identified citations will be screened based on inclusion criteria, collated, and uploaded into EndNote, and duplicates will be removed. Following a pilot test, titles and abstracts will be screened by two or more independent reviewers to assess the inclusion criteria. Potentially relevant studies will be retrieved in full, and their citation details will be imported into the JBI System for the Unified Management, Assessment, and Review of Information (JBI SUMARI) (JBI, Adelaide, Australia) [26]. The full text of selected citations will be assessed in detail against the inclusion criteria by two independent reviewers. The reasons for the exclusion of papers in the full text that do not meet the inclusion criteria will be recorded and reported in the systematic review. Any disagreements between the reviewers at any stage of the selection process will be resolved through discussion or consultation with a third reviewer. The search results and study inclusion process will be reported in full in the final systematic review and presented in a Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) flow diagram [27].

3.4 Assessment of methodological quality

The methodological quality of studies deemed eligible for inclusion will be assessed using the COSMIN checklist [28]. Two reviewers will conduct independent appraisals and, thereafter, achieve consensus on study ratings. A third reviewer will be consulted, as necessary. The COSMIN Risk of Bias checklist is modular and contains 10 boxes with standards for tool development (Box 1) and nine measurement properties: content validity (Box 2), structural validity (Box 3), internal consistency (Box 4), cross-cultural validity/measurement invariance (Box 5), reliability (Box 6), measurement error (Box 7), criterion validity (Box 8), hypothesis testing for construct validity (Box 9), and responsiveness (Box 10). This review will use specific standards for tool development (Box 1), content validity (Box 2), structural validity (Box 3), internal consistency (Box 4), and reliability (Box 6) because mixed methods appraisal tools are used by researchers to assess empirical study papers. Each criterion on the checklist will be rated as “very good,” “adequate,” “doubtful,” or “inadequate.” Studies will then receive an overall rating for methodological quality, which will be based on the criterion receiving the lowest rating—that is, the worst score will be considered the overall score. A working document created in an Excel file on the COSMIN website will be used to organize the appraisal results. The certainty of the findings will be assessed using GRADE—that is, Grading of Recommendations, Assessment, Development, and Evaluations [29].

3.5 Data extraction [28]

Data concerning the study characteristics, tool characteristics, and measurement properties of the data will be extracted. For characteristics, data regarding purpose (original intended use), application methods (checking methods, scoring methods, and answer format), administration burden (training necessity, time for completion), and number of items and topics will be collected. For measurement properties, validity and reliability results will be collected. The extracted information will be directly copied and pasted into tables (Appendix II). Further, data extraction will be performed by two independent reviewers to avoid missing relevant information. Next, the results of each measurement

property study will be evaluated according to the updated criteria for suitable measurement properties. Each result will be rated as sufficient (+), insufficient (-), or indeterminate (?). The results of each measurement property and its quality rating will also be directly added to the applicable table.

4. Data synthesis

The results will be quantitatively pooled and qualitatively summarized. These findings will be reported in a table with the rating of the pooled or summarized results and the rating of the quality of evidence (high, moderate, low, or very low). If possible, results from different studies on one measurement property will be statistically pooled in the meta-analysis. Findings that cannot be statistically pooled will be presented in a narrative form, and tables and figures will be included to aid data presentation, wherever appropriate.

5. Conclusion

In this systematic review protocol, the measurement properties of mixed methods appraisal tools are identified. The studies published and dissertations are searched from a variety of sources. Consensus on the appraisal tools of mixed methods studies will be benefit to ensure that such studies are properly planned, conducted, reported, and reviewed.

Compliance with ethical standards

Acknowledgments

The authors would like to thank the JBI comprehensive systematic review training program at the center of JBI-The University of Adelaide in July 2020. We appreciate assistance with the search strategy from Librarian Mr. Hiroshi Takahashi, Akita University. We would also like to thank Editage (www.editage.com) for English language editing. Reference manager software is EndNote V7.8 (Clarivate Analytics, PA, USA).

Disclosure of conflict of interest

The founders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Funding

This study was funded by the Japan Grant-in-Aid for Scientific Research (B) (20H03966).

References

- [1] Akagawa Y, Makabe S, Ito T, Kimura Y, Andoh H. Challenge and hope for parents who have cancer. *International Journal of Science and Research Archive*. 2021; 3(1): 136–147.
- [2] Fetter MD. *The mixed methods research workbook*. Los Angeles: SAGE; 2020.
- [3] Creswell JW, Plano Clark VL. *Designing and Conducting Mixed Methods Research*, Chapter 2, The foundation of mixed methods research. SAGE; 2018. p.21-49
- [4] Bazeley P. *Integrating Analyses in Mixed Methods Research*. SAGE; 2018. p.7
- [5] Fetter MD, Molina-Azorin JF. A Checklist of Mixed Methods Elements in a Submission for Advancing the Methodology of Mixed Methods Research. *Journal of Mixed Methods Research*. 2019; 13(4): 414-423.
- [6] O'Cathain A. Assessing the quality of mixed methods research: Toward a comprehensive framework. In A. Tashakkori & C. Teddlie (Eds.), *SAGE handbook of mixed methods in social and behavioral research*, (2nd ed.). SAGE; 2010. p.351-555.
- [7] Onwuegbuzie A, Burke RJ. The validity Issue in Mixed Research. *Research in the School*. 2006; 13(1): 48-63.
- [8] Pluye P, Gagnon MP, Griffiths F, Johnson-Lafleur J. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in Mixed Studies Reviews. *International Journal of Nursing Studies*. 2009 Apr; 46(4): 529-546.
- [9] Hong QN, Pluye P, Fábregues S, Bartlett G, Boadman F, Cargo M, Dagenais P, Gagnon M, Griffiths F, Nicolau B, O'cathain A, Rousseau M, Vedel I. MIXED METHODS APPRAISAL TOOL (MMAT). [Internet]. [cited 2021 Mar 8].

Available from:
http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/attach/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf.

- [10] Creamer EG. An introduction to fully integrated mixed methods research. USA: Blacksburg; 2018.
- [11] Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. SAGE; 2007.
- [12] O'Cathain A, Murphy E, Nicholl J. The quality of mixed methods studies in health services research. *Journal of Health Services Research & Policy*. 2008 Apr; 13(2): 92-98.
- [13] Fábregues S, Molina-Azorin JF. Addressing quality in mixed methods research: a review and recommendations for a future agenda. *Quality and Quantity*. 2017; 51: 2847–2863.
- [14] Heyvaert M, Hannes K, Maes B, Onghena P. Critical Appraisal of Mixed Methods Studies. *Journal of Mixed Methods Research*. 2013; 7 (4): 302-327.
- [15] Younas A, Pedersen M, Tayaben JL. Review of Mixed-Methods Research in Nursing. *Nurs Res*. 2019 Nov/Dec; 68(6): 464-472.
- [16] Fábregues S, Hong QN, Escalante-Barrios EL, Guetterman TC, Meneses J, Fetters MD. A Methodological Review of Mixed Methods Research in Palliative and End-of-Life Care (2014-2019). *International Journal of Environmental Research and Public Health*. 2020 May 29; 17(11).
- [17] Pace R, Pluye P, Bartlett G, Macaulay AC, Salsberg J, Jagosh J, Seller R. Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *International Journal of Nursing Studies*. 2012; 49(1): 47-53.
- [18] Hong QN, Gonzalez-Reyes A, Pluye P. Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies, the Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice*. 2018; 24(3): 459-467.
- [19] Hong QN, Pluye P, Fabregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, Gagnon MP, Griffiths F, Nicolau B, O'Cathain A, Rousseau MC, Vedel I. Improving the content validity of the mixed methods appraisal tool: a modified e-Delphi study. *Journal of Clinical Epidemiology*. 2019 Jul; 111: 49-59 e41.
- [20] COSMIN Guideline for Systematic Reviews of Outcome Measurement Instruments, 2021. [Internet]. [cited 2021 Sep 13] Available from: <https://www.cosmin.nl/tools/guideline-conducting-systematic-review-outcome-measures/>
- [21] Harrison R, Jones B, Gardner P, Lawton R. Quality assessment with diverse studies (QuADS): an appraisal tool for methodological and reporting quality in systematic reviews of mixed- or multi-method studies. *BMC Health Services Research*. 2021 Feb 15; 21(1): 144.
- [22] Headley MG, V.L. PC. Multilevel Mixed Methods Research Designs: Advancing a Refined Definition. *Journal of Mixed Methods Research*. 2019; 14(3): 145-163
- [23] Terwee CB, Prinsen CA, Chiarotto A, Vet Hcd, Bouter LM, Alonso J, Westerman MJ, Patrick DL, Mokkink LBao]. COSMIN methodology for assessing the content validity of PROMs. 2018. [Internet]. [cited 2022 Sep 12]. Available from: <https://cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf>
- [24] Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, Vet Hcd, et al. COSMIN methodology for systematic reviews of Patient-Reported Outcome Measures (PROMs) 2018. [Internet]. [cited 2022 Sep 12]. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf
- [25] Stephenson M, Riitano D, Wilson S, Leonardi-Bee J, Mabire C, Cooper K, Monteiro da Cruz D, Moreno-Casbas M, Lapkin S. Chapter 12: Systematic reviews of measurement properties. In: Aromataris E, Munn Z (Editors). *JBIM Manual for Evidence Synthesis*. JBI, 2020. [Internet]. [cited 2021 Mar 19]. Available from: <https://doi.org/10.46658/JBIMES-20-13>
- [26] Munn Z, Aromataris E, Tufanaru C, Stern C, Porritt K, Farrow J, Lockwood C, Stephenson M, Moola S, Lizarondo L, McArthur A, Peters M, Pearson A, Jordan Z. The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). *International Journal of Evidence-Based Healthcare*. 2019 Mar; 17(1): 36-43.
- [27] PRISMA. PRISMA TRANSPARENT REPORTING of SYSTEMATIC REVIEWS and META-ANALYSES [Internet]. [cited 2021 Mar 5]. Available from: <http://prisma-statement.org/>

- [28] Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, Vet HCd, Terwee CBao]. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. 2019. [Internet]. [cited 2022 Sep 12]. Available from: https://www.cosmin.nl/wp-content/uploads/COSMIN-study-designing-checklist_final.pdf#
- [29] GRADE working group. What is GRADE [Internet]. [cited 2022 April 8]. Available from: <https://www.gradeworkinggroup.org/>

Appendix

Appendix I Search strategy

For systematic reviews, all search strategies should be presented. Search conducted on Scopus at 28th of June 2022.

Search	Query	Records retrieved
#1	TITLE-ABS-KEY ("mixed methods")	67,689
#2	TITLE-ABS-KEY ("appraisal tool")	2,654
#3	TITLE-ABS-KEY (checklist)	84,450
#4	TITLE-ABS-KEY ("risk of bias assessment")	3,662
#5	TITLE-ABS-KEY (criteria)	1,709,980
#6	#2 OR #3 OR #4 OR #5	1,787,275
#7	TITLE-ABS-KEY (reliability)	991,442
#8	TITLE-ABS-KEY (validity)	617,340
#9	#7 OR #8	1,489,078
#10	#1 AND #6 AND #9	366

Appendix II: Data collection sheet

		Study 1	Study 2
Study characteristics	Author (year)		
	Title		
	Country		
	Sample size		
	Sample characteristics (age, gender, specialties)		
	Study purpose		
	Study design		
Tool characteristics	Development processes		
	Purpose (original intended use)		
	Application methods (checking methods, scoring methods, and answer format et al)		
	Administration burden (training necessity, time for completion)		
	Number of items and what topics.		
Measurement property data	Content validity		
	Structural validity		
	Internal consistency		
	Others		