

An aggregated quality assessment tool for literature review of disparate or homogenous data.

Ward,* Kim and Smith, Melody

Te Kura Naahi | School of Nursing, *Waipapa Taumata Rau* | The University of Auckland

[*k.ward@auckland.ac.nz](mailto:k.ward@auckland.ac.nz)

<https://unidirectory.auckland.ac.nz/profile/k-ward>

The systematised literature review is a systematic process using a prescribed protocol for identifying or screening literature; a valuable method for synthesising data from empirical studies to provide more substantial evidence related to a topic than the separate studies offer alone. Systematised reviews are a common method of knowledge synthesis for establishing current evidence on a wide variety of topics, from healthcare and nursing to education and engineering [1, 2]. Grant and Booth [3] identified up to 14 types of literature review, including critical review, mixed methods review, scoping review, state-of-the-art review, and umbrella review [4]. Literature reviews can synthesise evidence from quantitative studies, such as systematic reviews and meta-analyses. Or reviews can synthesise specifically qualitative data as with the meta-synthesis. However, other types of review, such as integrative, scoping and narrative reviews can incorporate disparate data from varied study types and from grey literature.

The process for conducting a systematised review of most types is well documented [3-11]. Typically, the process described for any style of review is robust and methodical [10]. Not all reviews require authors to quality assess included studies as part of the review process, such as with a critical review & mapping review. However, quality assessment is becoming more evident, particularly with integrative review [8]. Quality assessment of studies includes quality evaluation to determine the reliability, validity and relevance of included material [11]. Including poor quality material in a review may distort synthesis and excluding poor quality studies that meet criteria may similarly bias synthesis. Determining the quality of included articles is only meaningful when considered within the context of data analysis and the review question. Therefore, having a way to identify the quality of included studies makes the process of review and of the quality evaluation transparent and enhances the rigor, strength and reporting of the review.

Quality assessment of studies is simpler when included literature uses single methodologies (i.e., quantitative versus qualitative) or where research design is similar (e.g., a review of randomised controlled trials). However, reviews that include studies adopting a wide variety of methods create complexity and challenges to quality assessment [12]. Such reviews are important in order to capture the gamut of evidence on a particular topic, without prioritising knowledge gained from one methodology over the other [6]. Currently, there is no gold standard for scoring approaches used to assess literature quality in literature reviews [11-13]. However, there are many quality appraisal tools widely available. Toronto and Remington [11] identify over 100 quality appraisal tools, nine of which are commonly used in nursing research, some of which are designed to appraise disparate study designs.

While most of these tools offer sound modes of quality assessment in their content, they offer contrasting ways to measure and compare quality across differing methodologies, as noted by Ward et al. [14]. In some cases, different quality assessment tools are utilised for quantitative and qualitative research in the same review. By their nature, such tools are not directly transferable between methodologies. For example, the Critical Appraisal Skills Programme (CASP) [15] quality assessment tool is used for appraising qualitative studies, while the Hawker et al. [16] quality assessment tool is used for evaluating quantitative studies due to its flexibility across quantitative designs. For these reasons both tools were used for the Ward et al. integrative review. Papers were assessed for methodological rigour and relevance to the research question. Assessment criteria included the provision of a clear statement of research aim; appropriateness of the methodology to the research question; appropriate recruitment strategy, data collection and analysis; evidence of ethics and attention to bias. Studies were assessed as good, fair, poor or very poor. However, balancing reporting between the two tools was problematic. To enable a consistent evaluation of both qualitative and quantitative research in the Ward et al. review, KW developed an aggregated tool using Microsoft Excel and developed from the CASP [15] and Hawker et al. [16] tools. The aggregated tool is structured into ten domains that mirror the sections of a manuscript. This tool has since been used successfully by students and colleagues [14, 17-21].

Since 2014 as students and colleagues have used the aggregated tool, the authors have integrated other commonly used quality appraisal sources. These include integrating elements of the Johns Hopkins Research Evidence Appraisal Tool [22], the updated CASP [23] checklists, the Mixed Methods Appraisal Tool [24], the Cochrane Handbook of Systematic Reviews [25] and early Baxter and Eyles [26] work on qualitative studies. Aggregated in this way the tool provides appraisal guidance for quantitative, qualitative, and mixed-method analysis and reporting, as well as guidance for single versus multiple research evidence. In terms of utility, the authors also sought to avoid over-complicating the descriptors.

The aggregated tool offers a numerical scoring system that is the same across all assessment domains, with criteria descriptors provided across each of the domains for both quantitative and qualitative material. Studies evaluated via these criteria are scored between 1 and 4 for each of the ten domains adding up to a total score of 40 as the highest score or ten as the lowest: a score of 10 - 16 is very poor, 17 - 24 is poor, 25 - 32 is fair and 33 - 40 is of good quality. Evaluative criteria are designed to allow evaluation of any study type and are outlined in Table 1 below.

While the aggregated tool has face validity from practical use to date, evaluation and critique require greater use to provide substantial validity. To that end, the authors offer the Excel spreadsheet tool here for use and adaptation, asking only that this work is cited and acknowledged. We invite you to contact us to let us know how you found the aggregated tool, how you've adapted it if you have, and for what kind of review you have used it.

k.ward@auckland.ac.nz

Table 1: Aggregated tool domains and criteria

Domains: Article section and main questions:		1. Abstract and title: Did the authors provide a clear description of the study?	2. Introduction and aims: Was the background relevant and up to date and was there a clear statement of the research aims?	3. Method and data: Is the method appropriate and clearly explained?	4. Sampling: Was the sampling strategy appropriate to address the aims?	5. Data collection: Were the data collected in a way that addressed the research question?
	Good (4)	Structured abstract with full information and clear title.	Full but concise background to the discussion/study containing up-to- date literature review and highlighting gaps in knowledge. Clear statement of aim AND objectives including research questions.	Method is appropriate and described clearly (e.g., questionnaires included). QN: Allocation concealment (if multiple groups). Blinding of participants and researchers. Blinding of outcome assessment. Complete (or high level of completion) outcome data. QL: Prolonged engagement; persistent observation; multiple or participant researchers; peer examination; thick description of the audit process or audit products.	Details (age/sex/ethnicity/context) of who participated and how they were recruited. Why this group was targeted. The sample size was justified for the study. QN: Response rates shown and explained. Random sampling? Low or no participant attrition over time? Sample size calculations for primary outcome measure? QL: Sampling method described, e.g. purposeful, convenience, snowball etc.	Clear details of data collection method/s and recording. Method/s and setting for data collection justified. QN: Sample size calculations. Validity and reliability of tools measured QL: Congruence with methodology. Evidence of topic guide, or interview method etc. Any modification of methods explained. Form of data clear - voice recordings, video notes etc.
	Fair (3)	Abstract with most of the information.	Some background and literature review. Research questions outlined.	Method appropriate, description could be better. Data types described.	Sample size justified. Most information given, but some missing.	Most of the above information given, but some missing.
	Poor (2)	Inadequate abstract.	Some background but no aim/objectives/questions OR Aims/ objectives but inadequate background.	Questionable whether method is appropriate. Method described inadequately. Little description of data.	Sampling mentioned but few descriptive details.	Data collection mentioned, but few descriptive details.
	Very poor (1)	No abstract.	No mention of aims/objectives. No background or literature review.	No mention of method, AND/OR Method inappropriate, AND/OR No details of data.	No details of sample.	No details about data collection given.
Author / Year	Title	Score (1-4)	Score (1-4)	Score (1-4)	Score (1-4)	Score (1-4)
Smith 2092	Example	4	4	3	3	3

6. Data analysis: Was the description of the data analysis sufficiently detailed?	7. Ethics and bias: Have ethical issues been addressed, and was necessary ethical approval gained? Has the relationship between researchers and participants been adequately considered?	8. Results: Is there a clear statement of the findings?	9. Transferability or generalisability: Are the findings of this study transferable or generalisable to a wider population?	10. Implications and usefulness: How important are these findings to policy and practice?	
<p>Clear description of how analysis was done.</p> <p>QN: Reasons for tests selected hypothesis driven/ numbers add up/ statistical significance discussed. Controlled/stratified for confounders</p> <p>If the unit of allocation and the unit of analysis are different, was cluster analysis done?</p> <p>QL: Description of how themes or concepts derived. Respondent validation or triangulation. Peer debriefing, negative case analysis, referential adequacy, member checking?</p>	<p>Ethics: Where necessary issues of confidentiality, sensitivity, and consent were addressed. Researcher was reflexive and/or aware of own bias.</p> <p>QN: Bias: selection bias, performance bias, detection bias, attrition bias?</p> <p>QL: Personal ethnocentricity and biases made clear.</p>	<p>Findings explicit, easy to understand, and in logical progression. Tables, if present, are explained in text. Results relate directly to aims. Sufficient data are presented to support findings.</p> <p>QN: Avoids reporting bias - is a study protocol publicly available (e.g., published as an article or available from research website).</p> <p>QL: Thick description evidenced with data. Are credibility, transferability, dependability, confirmability of findings clear?</p>	<p>Context and setting of the study is described sufficiently to allow comparison with other contexts and settings, plus high score in Question 4 (sampling).</p> <p>QN: Individuals selected to participate in the study likely to be representative of the target population. Statement of limitations as to the generalisability of findings.</p> <p>QL: Statement of limitations as to the transferability of findings. Resonance of findings to other settings.</p>	<p>Contributes something new and/or different in terms of understanding / insight or perspective. Suggests ideas for further research. Suggests implications for policy and/or practice.</p>	
Qualitative: Descriptive discussion of analysis. Quantitative.	Lip service was paid to above (i.e., these issues were acknowledged).	Findings mentioned but more explanation could be given. Data presented relate directly to results.	Some context and setting described, but more needed to replicate or compare the study with others, PLUS fair score or higher in Question 4.	Two of the above (state what is missing in comments).	
Minimal details about analysis.	Brief mention of issues.	Findings presented haphazardly, not explained, and do not progress logically from results.	Minimal description of context/setting.	Only one of the above	
No discussion of analysis.	No mention of issues.	Findings not mentioned or do not relate to aims.	No description of context/setting.	None of the above	
Score (1-4)	Score (1-4)	Score (1-4)	Score (1-4)	Score (1-4)	TOTAL
4	4	4	3	3	35

References

- [1] Torraco RJ. Writing Integrative Reviews of the Literature: Methods and Purposes. *International journal of adult vocational education and technology*. 2016;7(3):62-70.
- [2] Perignat E, Katz-Buonincontro J. STEAM in practice and research: An integrative literature review. *Thinking Skills and Creativity*. 2019;31:31-43.
- [3] Grant MJ, Booth A. A typology of reviews: an analysis of 14 review types and associated methodologies. 2009;26(2):91-108.
- [4] Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research synthesis methods* [Internet]. 2014;5(4):371-85. <https://doi.org/10.1002/jrsm.1123>
- [5] Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol*. 2018;18(1):143.
- [6] Whittemore R, Knafl K. The integrative review: updated methodology. *J Adv Nurs* [Internet]. 2005;52(5):546-53. <https://doi.org/10.1111/j.1365-2648.2005.03621.x>
- [7] Gopalakrishnan S, Ganeshkumar P. Systematic Reviews and Meta-analysis: Understanding the Best Evidence in Primary Healthcare. *Journal of family medicine and primary care* [Internet]. 2013 PMC3894019];2(1):9-14. <https://doi.org/10.4103/2249-4863.109934>
- [8] Whittemore R, Chao A, Jang M, Minges KE, Park C. Methods for knowledge synthesis: An overview. *Heart & Lung: The Journal of Cardiopulmonary and Acute Care* [Internet]. 2014;43(5):453-61. <https://doi.org/10.1016/j.hrtlng.2014.05.014>
- [9] Dixon-Woods M, Sutton A, Shaw R, Miller T, Smith J, Young B, et al. Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. *J Health Serv Res Policy*. 2007;12(1):42-7.
- [10] Price C. Syntheses Synthesized: A Look Back at Grant and Booth's Review Typology. *Evidence Based Library and Information Practice* [Internet]. 2022;17(2):132-8. Epub Classics and Special Topics. <https://doi.org/10.18438/ebliip30093>
- [11] Toronto CE, Remington R. A Step-by-Step Guide to Conducting an Integrative Review [Internet]Switzerland Springer Cham; 2020 [Available from: <https://doi.org/10.1007/978-3-030-37504-1>
- [12] Whittemore R. Combining Evidence in Nursing Research: Methods and Implications. *Nurs Res* [Internet]. 2005;54 (1):56-62.
- [13] Quigley JM, Thompson JC, Halfpenny NJ, Scott DA. Critical appraisal of nonrandomized studies—A review of recommended and commonly used tools2019;25(1):44-52. <https://doi.org/10.1111/jep.12889>
- [14] Ward K, Hoare K, Gott M. What is known about the experiences of using CPAP for OSA from the users' perspective? A systematic, integrative literature review. *Sleep Med Rev*. 2014;18(4):357-66.
- [15] Critical Appraisal Skills Programme (CASP). Critical Appraisal Skills Programme (CASP): making sense of evidence about clinical effectiveness [Internet]Oxford: Minervation; 2010 [Available from: <http://www.casp-uk.net/find-appraise-act/appraising-the-evidence/>
- [16] Hawker S, Payne S, Kerr C, Hardey M, Powell J. Appraising the evidence: reviewing disparate data systematically. *Qual Health Res*. 2002;12(9):1284-99.
- [17] Watkins S, Isichei O, Gentles TL, Brown R, Percival T, Sadler L, et al. What is Known About Critical Congenital Heart Disease Diagnosis and Management Experiences from the Perspectives of Family and Healthcare Providers? A Systematic Integrative Literature Review. *Pediatric cardiology* [Internet]. 2022. <https://doi.org/10.1007/s00246-022-03006-8>

- [18] Browne T, Slark J, Ward K. How can you tell your nursing student has chosen the wrong vocation? An integrative literature review. *Nurse Educ Today*. 2020;93:104528.
- [19] Chaplin H, Ward K. How many hours per night is enough? A systematic integrative review to identify optimal hours of CPAP therapy use for sleep apnoea. *Health Sciences Review*. 2022;5:100061.
- [20] Ferguson A, Ward K, Parke R. Resilient healthcare and natural disasters: A scoping review *BMC Health Services Research* Under Review.
- [21] Jacob J. The course of post-operative pain following total hip or knee arthroplasty: The perception of orthopaedic patients in a New Zealand tertiary hospital. Unpublished: Te Kura Naahi, School of Nursing: The University of Auckland; 2019.
- [22] Dang D, Dearholt S. Johns Hopkins nursing evidence-based practice: Model & guidelines [Internet]. Third edition. ed. Indianapolis, IN: Sigma Theta Tau International; 2018 [Available from: https://www.hopkinsmedicine.org/evidence-based-practice/ijhn_2017_ebp.html]
- [23] Programme. CAS [Internet].2018. CASP Qualitative Checklist & Randomised Controlled Trial Checklist. Available from: <https://casp-uk.net/casp-tools-checklists/>
- [24] Hong QN, Fàbregues S, Bartlett G, Boardman F, Cargo M, Dagenais P, et al. The Mixed Methods Appraisal Tool (MMAT) version 2018 for information professionals and researchers. *Education for information*. 2018;34(4):285-91.
- [25] Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011; 2011.
- [26] Baxter JS, Eyles J. Evaluating Qualitative Research in Social Geography: Establishing 'Rigour' in Interview Analysis. *Transactions of the Institute of British Geographers*. 1997;22:505-25.