PROOF

Methodological quality of case series studies: an introduction to the JBI critical appraisal tool

Zachary Munn[1] • Timothy Hugh Barker[1] • Sandeep Moola[1,2] • Catalin Tufanaru[1,3] • Cindy Stern[1] • Alexa McArthur[1] • Matthew Stephenson[1] • Edoardo Aromataris[1]

[1]JBI, Faculty of Health and Medical Sciences, The University of Adelaide, Adelaide, Australia, [2]The George Institute for Global Health, Telangana, India, and [3]Australian Institute of Health Innovation, Faculty of Medicine and Health Sciences, North Ryde, Australia

Downloaded from https://journals.lww.com/jbisrir by BhDMf5ePHKav1zEoum1tQfN4a+kJLhEZgbsIHo4XMi0hCywCX1AWnYQp/IlQrHD3srV2EAJ191JKGzX8byPMeoq7KHNULVVn+w1pdkXE= on 11/04/2019

ABSTRACT

Introduction: Systematic reviews provide a rigorous synthesis of the best available evidence regarding a certain question. Where high-quality evidence is lacking, systematic reviewers may choose to rely on case series studies to provide information in relation to their question. However, to date there has been limited guidance on how to incorporate case series studies within systematic reviews assessing the effectiveness of an intervention, particularly with reference to assessing the methodological quality or risk of bias of these studies.

Methods: An international working group was formed to review the methodological literature regarding case series as a form of evidence for inclusion in systematic reviews. The group then developed a critical appraisal tool based on the epidemiological literature relating to bias within these studies. This was then piloted, reviewed and approved by the international Scientific Committee of JBI.

Results: The JBI critical appraisal tool for case series studies includes 10 questions addressing the internal validity and risk of bias of case series designs, particularly confounding, selection and information bias, in addition to the importance of clear reporting.

Conclusion: In certain situations, case series designs may represent the best available evidence to inform clinical practice. The JBI critical appraisal tool for case series offers systematic reviewers an approved method to assess the methodological quality of these studies.

Keywords Assessment of methodological quality; case series; methodological guidance; risk of bias; systematic review

JBI Database System Rev Implement Rep 2019; 17(0):1–7.

Introduction

Systematic reviews aim to provide a comprehensive and rigorous synthesis of the best available evidence to inform health policy and practice.[1-4] Although case series are not traditionally included in systematic reviews assessing the effectiveness of an intervention or therapy, they have contributed greatly to the medical literature and can offer valuable information relating to the benefits and harms of certain treatments.[5,6] As such, case series can be considered for inclusion in systematic reviews of effectiveness, particularly in the absence of experimental designs such as randomized controlled trials and observational analytical studies.

There is an element of confusion regarding both the nomenclature and characteristics of a case series, with the definition varying across the medical literature, resulting in the inconsistent use of the term.[7-9] The gamut of case series is wide, with some studies claiming to be a case series that are realistically a collection of case reports, while others are more akin to cohort studies or even quasi-experimental, before and after studies. This has created difficulty in assigning case series a position in the hierarchy of evidence and identifying an appropriate critical appraisal tool.[7-9] This is not only a challenge with case series studies but also within the broader epidemiological literature, resulting in efforts to classify

Correspondence: Timothy Hugh Barker, timothy.barker@adelaide.edu.au

The authors declare no conflict of interest.

DOI: 10.11124/JBISRIR-D-19-00099

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

and group various features of different types of research studies through the use of algorithms or flowcharts.[10-13] In these guidance documents, case series are described as an observational and non-comparative study design.

According to one dictionary of epidemiology, case series are "a collection of subjects (usually, patients) with common characteristics used to describe some clinical, pathophysiological, or operational aspect of a disease, treatment, exposure, or diagnostic procedure".[14(p.33)] It is noted that a case series "does not include a comparison group and is often based on prevalent cases and on a sample of convenience".[14(p.33)] In our view, case series are best described as observational (that is, not experimental and not randomized), descriptive studies, without a control (or comparator group). Dekkers *et al.*[8] define a case series as a study in which "only patients with the outcome are sampled (either those who have an exposure or those who are selected without regard to exposure), which does not permit calculation of an absolute risk".[(p.39)] The outcome could be a disease or disease-related. This is contrasted to cohort studies, where sampling is based on exposure (or characteristic), and case-control studies, where there is a comparison group without the disease.

All systematic reviews incorporate a process of critiquing and appraising the research evidence. The purpose of this appraisal is to assess the methodological quality of a study and to determine the extent to which a study has addressed the possibility of bias in its design, conduct and analysis.[15] All studies selected for inclusion in the systematic review (that is, those that fulfill the a priori eligibility criteria described in the protocol) need to be subjected to rigorous assessment of their quality of conduct by two independent critical appraisers. The results of this appraisal can then be used to inform synthesis and interpretation of the results of the systematic review.[15]

Systematic reviews often use critical appraisal tools that are study design specific. There may be separate tools used to appraise randomized controlled trials, cohort studies, cross-sectional studies and so on.[4] Because case series are an uncontrolled (and non-experimental) study design, they are associated with an increased risk of bias[5] and must be appraised with the same scrutiny expected of study designs associated with higher levels of evidence.[16] For example, the completeness of a case series contributes to its reliability,[8] with studies that indicate a consecutive and complete inclusion considered more reliable than those that do not.

The JBI approach to systematic reviews is one of pragmatism, where the aim is to include a summary of the best available evidence and not only randomized controlled trials.[17-19] As such, there was a need for a standardized tool that would allow for transparent and repeatable appraisal of case series included in systematic reviews of effectiveness. This paper documents the process of the creation and application of such a tool.

Methods

In 2014, a working group of researchers and methodologists was formed within JBI to investigate the use of case series studies in systematic reviews and the development of a critical appraisal tool for these designs. It was clear from the beginning that the group needed to ensure a clear understanding and definition for case series among all members. The group agreed with the principles outlined by Dekkers *et al.*[8] and defined case series as studies where only patients with a certain disease or disease-related outcome are sampled. Before proceeding to develop a new tool, the group conducted a search and review of existing methodological, epidemiological and health research literature on case series as well as previously published appraisal tools for case series. Although few guides and tools were identified,[20-22] the group felt that this guidance inadequately covered all important methodological areas specific to case series designs. Over a period of one year with many methodological discussions, the group developed a tool, which was then piloted internally by the authors. Items covered in the tools were selected based on our review of the methodological literature and relevant items from other JBI tools. Based on the results of this pilot, the final tool was then drafted and sent to the JBI international scientific committee for further review and feedback. Following minor modifications, the tool was approved by the Scientific Committee and made available to JBI reviewers (Table 1).[23] It was also embedded in the systematic review software JBI System for the Unified Management, Assessment and Review of Information (JBI SUMARI; Adelaide, Australia, JBI).[24]

PROOF

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

Table 1: JBI's tool for assessing case series

Question	Yes	No	Unclear	Not applicable
1. Were there clear criteria for inclusion in the case series?	○	○	○	○
2. Was the condition measured in a standard, reliable way for all participants included in the case series?	○	○	○	○
3. Were valid methods used for identification of the condition for all participants included in the case series?	○	○	○	○
4. Did the case series have consecutive inclusion of participants?	○	○	○	○
5. Did the case series have complete inclusion of participants?	○	○	○	○
6. Was there clear reporting of the demographics of the participants in the study?	○	○	○	○
7. Was there clear reporting of clinical information of the participants?	○	○	○	○
8. Were the outcomes or follow-up results of cases clearly reported?	○	○	○	○
9. Was there clear reporting of the presenting site(s)/clinic(s) demographic information?	○	○	○	○
10. Was statistical analysis appropriate?	○	○	○	○

Within the tool, some of the items relate to risk of bias, whereas others relate to ensuring adequate reporting and statistical analysis. A response of "no" to any of the following questions negatively impacts the overall quality of a case series.

How to use this tool

1. *Were there clear criteria for inclusion in the case series?*
   The authors should provide clear inclusion (and exclusion criteria where appropriate) for the study participants. The inclusion/exclusion criteria should be specified (e.g. risk, stage of disease progression) with sufficient detail and all the necessary information critical to the study.
2. *Was the condition measured in a standard, reliable way for all participants included in the case series?*
   The study should clearly describe the method of measurement of the condition. This should be done in a standard (i.e. same way for all patients) and reliable (i.e. repeatable and reproducible results) way.
3. *Were valid methods used for identification of the condition for all participants included in the case series?*
   Many health problems are not easily diagnosed or defined, and some measures may not be capable of including or excluding appropriate levels or stages of the health problem. If the outcomes were assessed based on existing definitions or diagnostic criteria, then the answer to this question is likely to be yes. If the outcomes were assessed using observer-reported or self-reported scales, the risk of over- or under-reporting is increased, and objectivity is compromised. Importantly, determine if the measurement tools used were validated instruments as this has a significant impact on outcome assessment validity.

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

4. *Did the case series have consecutive inclusion of participants?*
   Studies that indicate a consecutive inclusion are more reliable than those that do not. For example, a case series that states, ''We included all patients (24) with osteosarcoma who presented to our clinic between March 2005 and June 2006'' is more reliable than a study that simply states, ''We report a case series of 24 people with osteosarcoma.''
5. *Did the case series have complete inclusion of participants?*
   The completeness of a case series contributes to its reliability.[8] Studies that indicate a complete inclusion are more reliable than those that do not. A stated above, a case series that states, ''We included all patients (24) with osteosarcoma who presented to our clinic between March 2005 and June 2006'' is more reliable than a study that simply states, ''We report a case series of 24 people with osteosarcoma.''
6. *Was there clear reporting of the demographics of the participants included in the study?*
   The case series should clearly describe relevant participants' demographics such as the following information where relevant: participant's age, sex, education, geographic region, ethnicity and time period.
7. *Was there clear reporting of clinical information of the participants?*
   There should be clear reporting of clinical information of the participants such as the following information where relevant: disease status, comorbidities, stage of disease, previous interventions/treatment, results of diagnostic tests, etc.
8. *Were the outcomes or follow-up results of cases clearly reported?*
   The results of any intervention or treatment should be clearly reported in the case series. A good case series should clearly describe the clinical condition post-intervention in terms of the presence or lack of symptoms. The outcomes of management/treatment when presented as images or figures can help in conveying the information to the reader/clinician. It is important that adverse events are clearly documented and described, particularly when a new or unique condition is being treated or when a new drug or treatment is used. In addition, unanticipated events, if any that may yield new or useful information should be identified and clearly described.
9. *Was there clear reporting of the presenting site(s)/clinic(s) demographic information?*
   Certain diseases or conditions vary in prevalence across different geographic regions and populations (e.g. women men, sociodemographic variables between countries). The study sample should be described in sufficient detail so that other researchers can determine if it is comparable to the population of interest to them.
10. *Was statistical analysis appropriate?*
    As with any consideration of statistical analysis, consideration should be given to whether there was a more appropriate alternate statistical method that could have been used. The methods section of studies should be detailed enough for reviewers to identify which analytical techniques were used and whether these were suitable.

Discussion

This critical appraisal tool for case series studies has been publicly available since December 2017, published on the JBI website. Since then, it has been appropriately used in systematic reviews[25-27] and also cited in systematic review protocols.[28-30] As evidenced from these reviews and protocols, the inclusion of a case series is typically most beneficial in reviews of effectiveness; prevalence and/or incidence; and etiology and/or risk, particularly when there are no other studies to consider.[31]

Recently, multiple new case series tools have been published in the literature.[5,22] Both Murad *et al.*[5] and Guo *et al.*[22] have documented the creation and provision of similar tools to evaluate the methodological quality of a case series. The majority of the questions included in both of these tools address similar issues as those presented in the JBI tool, with minor variations in wording that could have ramifications for how appraisers interpret the results after using each tool. While there are advantages and disadvantages associated with each tool, the assessment of risk of bias, particularly when assessing observational studies, may be too complex a task for any single tool.[32] Assessment of risk of bias can be further hampered by a lack of compatibility between the chosen tool and the review team, with

PROOF

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

some teams feeling more comfortable or familiar with one tool over another. This highlights the importance of piloting the tool during the critical appraisal process. Some questions may need to be tailored to suit the research focus, while a scoring framework of ''yes,'' ''no'' or ''not applicable (N/A)'' may not be suitable for each question relevant to the study design or the research parameters. It is our opinion that all three tools provide a clear and logical format for the appraisal of case series. The advantage of the Guo *et al.*[22] and JBI tools is that they are designed specifically for case series (as opposed to a joint tool for case series and case reports), which allows them to have additional questions and be more specific to case series designs.[32]

In addition to these tools, there is a tool that has been developed to assess the risk of bias when conducting an effectiveness review and when including non-randomized studies: the Risk of Bias in Non-randomised Studies of Interventions (ROBINS-I).[33] This is a domain-based tool, and although it is designed for non-randomized studies, it is particularly designed for studies with ''cohort-like'' designs or designs with a control group. As case series studies do not have a control group, this tool may not be ideal for these study designs, nor would the Newcastle-Ottawa scale.[34] In a review of tools for critically appraising non-randomized studies, Quigley *et al.*[35] recommend that because there is no current consensus on which tool to use, systematic reviewers should select an appropriate tool based on the study design of selected papers for inclusion in their review.

Many researchers prefer using a domain-based approach in the critical appraisal of primary literature.[36,37] The main domains of bias assessed in observational studies include confounding bias, selection bias and information bias (including measurement, detection, classification, analysis and reporting bias), and these biases can be assessed through the use of signaling questions.[37] Although the tool presented in this paper has not been designed based on the domain approach, the questions can be seen as signaling questions for particular domains of bias. For example, questions 1, 4 and 5 can be considered signaling questions for the domain ''bias in selection of participants into the study''; questions 2 and 3 for the domain ''bias in measurement of outcomes''; questions 6 and 7 for the domain ''bias in selection of the reported results''; and question 8 for the domain ''bias due to missing data''.

We often receive queries from reviewers wishing to use our tool to provide advice on how much weight to assign each question, and what the cut-off score should be for inclusion in a systematic review. These questions presuppose that the purpose of appraisal in systematic reviews is to include only those studies that are of high quality and to exclude those of poor quality. While this is one way to use the results of critical appraisal in reviews, it is not the only approach, and it may not be appropriate in many situations. Our guidance to authors wishing to use this tool in terms of cut-off values/scores and determining whether a study is low, moderate or high quality, is that these thresholds are best decided by the systematic reviewers themselves. Generally, cut-off scores are advised against, because the critical appraisal questions are not all ''equal.'' As such, simply tallying the ''yes'' responses does not truly give an accurate indication of the specific problems of a study. We suggest presenting the results of critical appraisal for all questions via a table rather than summarizing with a score. Ideally, two reviewers will be involved in the critical appraisal for the review.

This tool is now in active use by the Joanna Briggs Collaboration and other systematic reviewers who are including case series designs in their reviews. This tool was developed based on methodological and epidemiological principles and has been reviewed internally by the author group in addition to the JBI Scientific Committee. It has been deemed as acceptable and appropriate by these groups and as such has demonstrated face validity. Further validation efforts are now required to establish the psychometric properties of the tool in addition to other issues, such as its acceptability, timeliness and ease of use. As previously described, another program of work relates to transferring this tool into a domain-based approach. However, given the lack of any current validated tool for critiquing case series studies, it was the view of the group that this tool be made widely available to assist systematic reviewers with the conduct of their reviews.

Conclusion

When there is limited availability of high-quality experimental studies (such as randomized controlled trials) on the effectiveness or harms of an intervention, case series may represent the best available evidence to inform clinical practice. As such, a

PROOF

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

critical appraisal tool is required. The JBI critical appraisal tool for case series offers systematic reviewers an approved method to assess the methodological quality of these studies.

Acknowledgments

The authors acknowledge the contribution of Dr. Kylie Porritt during the drafting of this paper.

References

1. Aromataris E, Pearson A. The systematic review: an overview. Am J Nurs 2014;114(3):53–8.
2. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. BMC Med Res Methodol 2018;18(1):5.
3. Munn Z, Tufanaru C, Aromataris E. JBI's systematic reviews: data extraction and synthesis. Am J Nurs 2014;114(7):49–54.
4. Porritt K, Gomersall J, Lockwood C. JBI's systematic reviews: study selection and critical appraisal. Am J Nurs 2014;114(6):47–52.
5. Murad MH, Sultan S, Haffar S, Bazerbachi F. Methodological quality and synthesis of case series and case reports. BMJ Evid Based Med 2018;23(2):60–3.
6. Vandenbroucke JP. In defense of case reports and case series. Ann Intern Med 2001;134(4):330–4.
7. Abu-Zidan F, Abbas A, Hefny A. Clinical "case series": a concept analysis. Afr Health Sci 2012;12(4):557–62.
8. Dekkers OM, Egger M, Altman DG, Vandenbroucke JP. Distinguishing case series from cohort studies. Ann Intern Med 2012;156(1 Part 1):37–40.
9. Esene IN, Ngu J, El Zoghby M, Solaroglu I, Sikod AM, Kotb A, *et al*. Case series and descriptive cohort studies in neurosurgery: the confusion and solution. Childs Nerv Syst 2014;30(8):1321–32.
10. Hartling L, Bond K, Harvey K, Santaguida PL, Viswanathan M, Dryden DM. Developing and testing a tool for the classification of study designs in systematic reviews of interventions and exposures [Internet]. Rockville, MD: Agency for Healthcare Research and Quality; 2010 ; [cited April 19, 2019]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK52670/.
11. Hartling L, Bond K, Santaguida PL, Viswanathan M, Dryden DM. Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy. J Clin Epidemiol 2011;64(8):861–71.
12. Peinemann F, Kleijnen J. Development of an algorithm to provide awareness in choosing study designs for inclusion in systematic reviews of healthcare interventions: a method study. BMJ Open 2015;5(8):e007540.
13. Seo H-J, Kim SY, Lee YJ, Jang B-H, Park J-E, Sheen S-S, *et al*. A newly developed tool for classifying study designs in systematic reviews of interventions and exposures showed substantial reliability and validity. J Clin Epidemiol 2016;70:200–5.
14. Porta M, editor. A dictionary of epidemiology. *A dictionary of epidemiology*. Oxford: Oxford University Press, 2014.
15. Averis A, Pearson A. Filling the gaps: identifying nursing research priorities through the analysis of completed systematic reviews. JBI Reports 2003;1(3):49–126.
16. The Joanna Briggs Institute Levels of Evidence and Grades of Recommendation Working Party. Supporting document for the Joanna Briggs Institute Levels of Evidence and Grades of Recommendation. 2014 [cited April 19, 2019]. Available from: https://joannabriggs.org/sites/default/files/2019-05/JBI%20Levels%20of%20Evidence%20Supporting%20Documents-v2.pdf.
17. Hannes K, Lockwood C. Pragmatism as the philosophical foundation for the Joanna Briggs meta-aggregative approach to qualitative evidence synthesis. J Adv Nurs 2011;67(7):1632–42.
18. Jordan Z, Lockwood C, Munn Z, Aromataris E. The updated Joanna Briggs Institute Model of Evidence-Based Healthcare. Int J Evid Based Healthc 2019;17(1):58–71.
19. Munn Z. Implications for Practice: should recommendations be recommended in systematic reviews? JBI Database System Rev Implement Rep 2015;13(7):1–3.
20. Pierson DJ. How to read a case report (or teaching case of the month). Respir Care 2009;54(10):1372–8.
21. Chan K, Bhandari M. Three-minute critical appraisal of a case series article. Indian J Orthop 2011;45(2):103.
22. Guo B, Moga C, Harstall C, Schopflocher D. A principal component analysis is conducted for a case series quality appraisal checklist. J Clin Epidemiol 2016;69:199–207.
23. Moola S, Munn Z, Tufanaru C, Aromataris E, Sears K, Sfetcu R, *et al*. Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E, Munn Z (Editors). Joanna Briggs Institute Reviewer's Manual [Internet]. Adelaide: Joanna Briggs Institute, 2017 [cited April 20, 2019]. Available from: https://reviewersmanual.joannabriggs.org/.
24. Munn Z, Aromataris E, Tufanaru C, Stern C, Porritt K, Farrow J, *et al*. 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). Int J Evid Based Healthc 2019;17(1):36–43.
25. Harris ES, Meiselman HJ, Moriarty PM, Metzger A, Malkovsky M. Therapeutic plasma exchange for the treatment of systemic sclerosis: A comprehensive review and analysis. J Scleroderma Relat Disord 2018;3(2):132–52.
26. Kuperus JS, Waalwijk JF, Regan EA, van der Horst-Bruinsma IE, Oner FC, de Jong PA, *et al*. Simultaneous occurrence of ankylosing spondylitis and diffuse idiopathic skeletal hyperostosis: a systematic review. Rheumatology 2018;57(12):2120–8.

PROOF

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.

27. Zucchelli G, Tavelli L, Ravidà A, Stefanini M, Suárez-López del Amo F, Wang HL. Influence of tooth location on coronally advanced flap procedures for root coverage. J Periodontol 2018;89(12):1428–41.
28. Di Castro VC, Hernandes JC, Mendonça ME, Porto CC. Life satisfaction and positive and negative feelings of workers: a systematic review protocol. Syst Rev 2018;7(1):243.
29. Eardley-Harris N, Munn Z, Cundy PJ, Gieroba TJ. The effectiveness of selective thoracic fusion for treating adolescent idiopathic scoliosis: a systematic review protocol. JBI Database System Rev Implement Rep 2015;13(11):4–16.
30. Ravat S, Olivier B, Gillion N, Lewis F. Laterality judgment performance between people with chronic pain and pain-free individuals: a systematic review protocol. JBI Database System Rev Implement Rep 2018;16(8):1621–7.
31. Fitzpatrick-Lewis D, Thomas H, Ciliska D. The methods for the synthesis of studies without control groups. Hamilton, ON: National Collaborating Centre for Methods and Tools; 2009.
32. Bero L, Chartres N, Diong J, Fabbri A, Ghersi D, Lam J, *et al*. The risk of bias in observational studies of exposures (ROBINS-E) tool: concerns arising from application to observational studies of exposures. Syst Rev 2018;7(1):242.
33. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, *et al*. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.
34. Wells G, Shea B, O'Connell D, Peterson J, Welch V, Losos M, *et al*. The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomized studies in meta-analysis. Ottawa, ON: The Ottawa Health Research Institute; 2011.
35. Quigley JM, Thompson JC, Halfpenny NJ, Scott DA. Critical appraisal of nonrandomized studies—A review of recommended and commonly used tools. J Eval Clin Pract 2019;25(1):44–52.
36. Bazerbachi F, Haffar S, Hussain MT, Vargas EJ, Watt KD, Murad MH, *et al*. Systematic review of acute pancreatitis associated with interferon-α or pegylated interferon-α: Possible or definitive causation? Pancreatology 2018;18(7):691–9.
37. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al*. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928.

PROOF

©2019 Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.