Mini Review Article   **Open Access**

Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis

Laleh Jamshidi[1]*, Lies Declercq[1], John M. Ferron[2], Mariola Moeyaert[3], S. Natasha Beretvas[4], and Wim Van den Noortgate[1]

[1] Faculty of Psychology and Educational Sciences & imec-Itec, KU Leuven (University of Leuven), Belgium
[2] University of South Florida, Tampa, Florida, USA
[3] University at Albany – State University of New York, New York, USA
[4] University of Texas at Austin, Texas, USA

ABSTRACT

Single-case experimental design (SCED) studies are becoming more prevalent in a variety of different fields and are increasingly included in meta-analyses (MAs) and systematic reviews (SRs). As MA/SR's conclusions are used as an evidence base for making decisions in practice and policy, the methodological quality and reporting standards of SRs/MAs are of uttermost importance. One way to improve the reliability and validity of SCED MAs and therefore to provide more confidence in MA/SR findings to practitioners and clinicians to decide on a particular intervention is the use of high-quality standards when conducting and reporting MAs/SRs. In the current study, some existing tools for assessing the quality of SRs/MAs that might also be helpful for SCED MAs will be reviewed briefly. These tools and guidelines can help meta-analysts, reviewers, and users to organize and evaluate the quality and reliability of the findings.

In order to investigate a certain intervention effect, the classic research design is a group comparison experimental design. In this kind of designs, the participants are randomly assigned to either intervention or control groups and the means of one or more dependent variables are compared to assess the effectiveness of the intervention. In order to get reliable effect size estimates and reach an acceptable level of statistical power, a large sample size of study participants is required in these designs. Single-case experimental designs (SCEDs) are alternative research designs that do not require many participants (or cases) and therefore are well suited to be used for studying rare phenomena, e.g., specific diseases or disabilities[1-3]. In this kind of designs, outcomes of interest are measured repeatedly for one or multiple cases under at least two conditions (i.e., typically a control phase followed by an intervention phase). Within each specific case, the measurements are compared across conditions or phases to investigate whether introducing the intervention has a causal effect on one or more outcomes[2,4-7]. SCEDs are frequently used in a variety of different fields such as psychology and educational sciences to evaluate the effectiveness of interventions of interest[7-11].

Due to the small number of participants, the main issue of SCEDs is limited generalizability of their findings. To overcome this issue of generalizability, SCEDs can be replicated across participants, and systematic review (SR) approaches can be applied to synthesize the results[4,12,13]. A SR is a kind of literature review to identify, evaluate, and aggregate all relevant studies on the same topic. In order to

decrease the possible systematic bias to answer particular research question(s), specific methods could be applied in SR[14]. A SR can include a meta-analysis (MA), which refers to a statistical integration of the findings from individual studies, typically by combining and comparing observed effect sizes[15].

SCED data have specific features that should be taken into account while calculating effect sizes in individual studies and synthesizing the effect sizes in a meta-analysis afterwards; otherwise, biased estimates might be obtained and statistical inferences may be flawed. For instance, the outcome variable could systematically decrease or increase over time even without being exposed to any intervention. Such a time trend should be accounted for in calculating effect sizes[4,16]. Another feature that has to be considered is the possible presence of serial dependency or autocorrelation in which the sequential measurements are more similar compared to farther measurements, violating the assumption of independence[17,18].

Conducting a SCED SR or MA could provide better insights into the overall effectiveness of interventions, as well as about factors that moderate the effect. Yet, poorly conducted SRs/MAs can lead to inaccurate inferences about the intervention effectiveness. Conclusions may be affected by deficiencies in designing, performing, and reporting these SRs/MAs. Therefore, it is important that users of SRs/MAs results (e.g., clinicians, researchers, and policy makers) consider the methodological quality of these studies. One way to do this is by assessing their quality by means of a standardized tool. Such a tool may also be useful for meta-analysts and systematic reviewers to ensure that their studies are well designed, conducted, and reported. On top of giving insight into the specific strengths and weaknesses of a study, such a tool can also be useful to assess the quality in general, although there is a considerable debate over using a quantifiable summary score to assess and rate the quality. The results of our recent systematic review of 178 SCED MAs conducted between 1985 and 2015[19] indicate that according to the R-AMSTAR, a considerable percentage of studies scored low on methodological quality. This tool assesses the methodological quality based on 11 main items that are further operationalized by means of 44 criteria. In order to apply the scale to SCED MAs rather than to MAs of group comparison studies, we had to reformulate some of the criteria. The MAs scored relatively high regarding some aspects such as "providing the characteristics of the included studies" and "doing a comprehensive literature search". The main deficiencies were related to "reporting an assessment of the likelihood of publication bias" and "using the methods appropriately to combine the findings of studies". In that review of SCED MAs, the methodological quality was evaluated by applying the modified R-AMSTAR, but there are other tools available that can be used. In the review of

Jamshidi et al. (in press)[19], the R-AMSTAR was chosen because it was found more comprehensive and detailed compared to other tools and due to its ability to produce a quantifiable assessment of methodological quality. More details related to the choice of the R-AMSTAR and the modified items can be found in that paper. In the current review, we give an overview of some of the frequently used tools for either assessing or improving the quality of SRs/MAs, and discuss their appropriateness for SCED SRs/MAs. To the best of our knowledge, there is no specific validated tool to assess the quality of SCED MAs or SRs and further research to produce a validated tool would be quite beneficial.

Approaches to Evaluate and Improve the Quality of Systematic Reviews and Meta-Analyses

To avoid inaccurate conclusions that might mislead decision-makers, meta-analysts and systematic reviewers should try to decrease key methodological deficiencies[20–22], such as not applying a random-effects model in case of heterogeneity, not assessing the likelihood of publication bias, or not assessing the scientific quality of included studies in formulating the conclusions, among others. Such kinds of deficiencies could also be expected to occur in SCED MAs and SRs. Conflicting results from SRs may confuse readers[23], and make it more difficult for practitioners and clinicians to make appropriate inferences. In order for systematic reviews and meta-analyses to provide valid and reliable evidence for informing decisions in research and policy-making, these must strictly uphold high methodological standards[21,23–25].

In addition, the users of SRs and MAs have a responsibility[26]: scientists, practitioners and clinicians should critically examine the methodological quality of a SR to avoid potentially misleading information when developing clinical decisions and guidelines[20,25,27,28].

Several tools have been developed specifically to assess the quality of SRs and MAs by either those who are conducting MA/SR or also those who use the results of MAs and SRs, such as practitioners and clinicians. By applying such tools, meta-analysts could ensure their studies meet high standards of quality, while users could be more informed on the reliability of MA or SR when basing their decisions. Table 1 lists some of the more well-known and commonly used tools[28–30], which have been specifically developed for assessing the quality of SRs/MAs and describes the basic features and guidance for their use (e.g. the purpose of the tool, the number of items, the items, and the judgement). Note that these tools are not specifically intended for meta-analyzing results from SCED studies. However, facets of these tools are useful and appropriate for judging the quality of SRs and MAs of SCED research studies.

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

Table 1: *Overview of characteristics of tools for assessing the quality of SRs/MAs*

Tool	Purpose	Items	Domains and items	Judgement

Sack's Quality Assessment Checklist [32]	Evaluating the quality of MAs	6 main domains and 23 items	Prospective design
			    Literature search
			    List of trials analyzed
			    Log of rejected trials
			    Treatment assignments
			    Ranges of patients
			    Ranges of diagnosis
			Combinability
			    Criteria
			    Measurement
			Control of bias
			    Selection bias
			    Data-extraction bias
			    Inter-observer agreement
			    Source of support
			Statistical analysis
			    Statistical methods
			    Statistical errors
			    Confidence intervals
			    Subgroup analyses
			Sensitivity analysis
			    Quality assessment
			    Varying methods
			    Publication bias
			Application of results
			    Caveats
			    Economic impact
				For each of these areas the checklist evaluates whether they are addressed in the systematic review or not. "Adequate" when the item had been fully addressed, "Partial" when some aspect was missing, and "None or unknown" when the item was not addressed

Overview Quality Assessment Questionnaire (OQAQ)[33]	Assessing the scientific quality of research overviews	10 items (9 individual items for assessing the quality and the last item is the overall assessment based on the first 9 items)	1. Were the research methods reported? 2. Was the search comprehensive? 3. Were the inclusion criteria reported? 4. Was selection bias avoided? 5. Were the validity criteria reported? 6. Was validity assessed appropriately? 7. Were the methods used to combine studies reported? 8. Were the findings combined appropriately? 9. Were the conclusions supported by the reported data? 10. What was the overall scientific quality of the overview?	Assessing a review's validity in terms of process rather than outcome. This tool can evaluate the potential threats to validity of this process. Clearly meets the criterion (scored as "yes"), clearly does not meet the criterion (scored as "no"), partially meet or is unclear whether it has met the criterion (scored as "partially") [34]
Assessment of Multiple Systematic Reviews (AMSTAR)[28]	Assessing the methodological quality of SRs	11 domains	1. Was an 'a priori' design provided? 2. Was there duplicate study selection and data extraction? 3. Was a comprehensive literature search performed? 4. Was the status of publication (i.e. grey literature) used as an inclusion criterion? 5. Was a list of studies (included and excluded) provided? 6. Were the characteristics of the included studies provided? 7. Was the scientific quality of the included studies assessed and documented? 8. Was the scientific quality of the included studies used appropriately in formulating conclusions? 9. Were the methods used to combine the findings of studies appropriate? 10. Was the likelihood of publication bias assessed? 11. Was the conflict of interest included?	Each individual item should be scored as one of the answers of "Yes", "No", "Can't answer", or "Not applicable"

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

Revised AMSTAR (R-AMSTAR) [35]	Assessing the methodological quality of SRs	11 main domains (from AMSTAR), operationalized with 41 criteria	Domain 1 (3 criteria) Domain 2 (3 criteria) Domain 3 (5 criteria) Domain 4 (4 criteria) Domain 5 (4 criteria) Domain 6 (3 criteria) Domain 7 (4 criteria) Domain 8 (4 criteria) Domain 9 (5 criteria) Domain 10 (3 criteria) Domain 11 (3 criteria)	Assessing the methodological quality of SRs in a quantifiable way. Each domain's score ranges from 1 to 4 (based on how many criteria were met) and the total score of R-AMSTAR would be calculated by summing the scores of all 11 domains and ranges from 11 to 44

Scottish Intercollegiate Guidelines Network (SIGN) [36]	Methodology checklist for SRs and MAs	12 domains	The same domains as AMSTAR, but domain 1 changed as follows:
			    The research question is clearly defined and the inclusion/ exclusion criteria must be listed in the paper.
			Domain 2 was divided in two separate domains as follows:
			    At least two people should have selected studies.
			    At least two people should have extracted data.
				Most of the items are scored with "yes" and "no". Other items have options such as "can't say" or/and "not applicable". The overall assessment of the study is judged as "high quality", "acceptable", "low quality", or "unacceptable".

Quality of Reporting of Meta-analyses (QUOROM) [37]	Checklist of standards for reporting the abstract, introduction, methods, results, and discussion sections of a meta-analysis	21 headings and subheadings	Title
			    Title
			Abstract
			    Objectives
			    Data sources
			    Review methods
			    Results
			    Conclusion
			Introduction
			Methods
			    Searching
			    Selection
			    Validity assessment
			    Data abstraction
			    Study characteristics
			    Quantitative data synthesis
			Results
			    Trial flow
			    Study characteristics
			    Quantitative data synthesis
			Discussion
				These standards encourage researchers to provide more information to readers about methods regarding the searches, study selection, validity assessment (e.g. quality assessment), data abstraction, study characteristics, and quantitative data synthesis, and about the results with regard to the 'trial flow', study characteristics, and quantitative data synthesis. They are also asked to provide information related to the number of trials identified, included, and excluded and to the reasons for exclusion

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) [38]	Checklist to report MAs and SRs	27 items in 7 main domains (adopted and extended version of QUOROM)	Title
			    Title
			Abstract
			    Structured summary
			Introduction
			    Rationale
			    Objectives
			Methods
			    Protocol and registration
			    Eligibility criteria
			    Information sources
			    Search
			    Study selection
			    Data collection process
			    Data items
			    Risk of bias in individual studies
			    Summary measures
			    Synthesis of results
			    Risk of bias across studies
			    Additional analyses
			Results
			    Study selection
			    Study characteristics
			    Risk of bias within studies
			    Results of individual studies
			    Synthesis of results
			    Risk of bias across studies
			    Additional analysis
			Discussion
			    Summary of evidence
			    Limitations
			    Conclusions
			Funding
			    Funding
				The tool aims to help researchers improve reporting of SRs and MAs of randomized trials as well as the types of experimental designs

Critical Appraisal Skills Programme (CASP) [39]	Appraise a SR	10 domains	1. Did the review address a clearly focused question? 2. Did the authors look for the right type of papers? 3. Do you think all the important, relevant studies were included? 4. Did the review's authors do enough to assess quality of the included studies? 5. If the results of the review have been combined, was it reasonable to do so? 6. What are the overall results of the review? 7. How precise are the results? 8. Can the results be applied to the local population? 9. Were all important outcomes considered? 10. Are the benefits worth the harms and costs?	Most of the questions are scored as "Yes", "No", and "Can't tell"
National Institute for Health and Care Excellence (NICE) [40]	Methodological checklist to assess the suitability of SRs and MAs to answer a guidance review question	5 domains	1. The review addresses an appropriate and clearly focused question that is relevant to the review question 2. The review collects the type of studies you consider relevant to the guidance review question 3. The literature search is sufficiently rigorous to identify all the relevant studies 4. Study quality is assessed and reported 5. An adequate description of the methodology used is included, and the methods used are appropriate to the question	The items are scored as "Yes", "No", and "Unclear"

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

Method-ological Expectations of Cochrane Intervention Reviews (ME-CIR)[41]	Standards for the reporting of new Cochrane Intervention Reviews	16 main domains with 109 standards	1. Title and authors (2 items) 2. Abstract (16 items) 3. Background (7 items) 4. Methods (7 items) 5. Search methods for identification of studies (6 items) 6. Data collection and analysis (17 items) 7. Description of studies (17 items) 8. Risk of bias in included studies (3 items) 9. Effects of interventions (24 items) 10. Discussion (2 items) 11. Authors conclusions (2 items) 12. Acknowledgement (1 item) 13. Contribution of authors (1 item) 14. Decleration of interest (1 item) 15. Differences between protocol and review (2 items) 16. Sources of support (1 item)	Very detailed and comprehensive standards. Most of the standards are mandatory to report (73%) and the rest of them are highly recommended

Risk of bias (ROBIS) tool[42]	Assess the risk of bias in a SR	4 main domains with 21 items	Study eligibility criteria
			    Five criteria, e.g., on clarity, relevance and the reflection of objectives, eligibility criteria, and restrictions on inclusion
			Identification and selection of studies
			    Five criteria, e.g., on search strategy, searching in databases or any additional methods, restrictions for search and selection, and minimizing the errors in selection
			Data collection and study appraisal
			    Five criteria, e.g., on minimizing the error in data collection, providing study characterisitics, collecting study results for synthesizing, assessing quality, and minimizing the risk of bias in assessment
			Synthesis and findings
			    Six criteria, e.g., on synthesizing all the studies, following all the predefined analyses, addressing heterogeneity, checking sensitivity analysis, checking or addressing the biases in primary studies.
				The items are scored as "Yes", "Probably Yes", "Probably No", "No" and "No Information", with "Yes" indicating low concerns. The subsequent level of concern about bias associated with each domain is then judged as "low," "high," or "unclear."

Some of these tools focus on the description of the methodology and findings (e.g. PRISMA and QUOROM) and some others concentrate on methodological quality and evaluate how well the SR was designed and performed (e.g. AMSTAR, R-AMSTAR, OQAQ)[31]. Some of the above-mentioned tools explicitly address that they can be used not only for conducting and reporting the MAs/SRs, but also for critical appraising published MAs/ SRs (e.g. Sack's checklist, PRISMA, QUOROM). Although it was not explicitly stated in descriptions of other tools whether they could be applied by meta-analysts and reviewers while they are performing and reporting their studies, we believe that being aware of criteria that might be used for critically appraising the quality of SRs/MAs could be helpful for researchers for designing, conducting, and reporting the results and conclusions of SRs/MAs.

Table 2 gives a further comparison of the content of the reviewed tools. Some tools assess one aspect of methodological quality via one general item, whereas others use multiple detailed criteria. The comparison indicates that providing search strategy, validity/ quality assessment of primary studies, and checking the possibility of combining the results are the aspects that were considered in all reviewed tools.

Discussion

SRs and MAs are essential methods for aggregating the results of primary studies in a specific field. Nevertheless, the reliability and validity of their associated conclusions could be compromised by the risk of methodological flaws. Since limited generalizability is a key limitation of SCED studies in providing a source of information for practitioners and clinicians to make the best decisions and guidelines for practice, conducting high-quality SCED MAs/SRs is of the uttermost importance. The recent results of the review of methodological quality of SCED MAs[19], indicate that improving the scientific quality of SCED MAs/SRs is necessary. Applying a validated tool (or using a modified tool or a combination of tools) consisting of methodological standards might be helpful for supporting meta-analysts/reviewers who are conducting studies

Journal of Mental Health & Clinical Psychology

Table 2: *Comparison of aspects related to methodological quality of SRs/MAs among tools*

Characteristics	Sack's Quality Assessment Checklist	OQAQ	AMSTAR	R-AMSTAR	SIGN	QUOROM	PRISMA	CASP	NICE	MECIR	ROBIS
Registered protocol	-	-			-	-		-	-		-
Addressing an appropriate and clearly focused questions	-	-									
Research methods/design					-	-	-	-	-		-
Search strategy											
Any restrictions for search (e.g. publication status, language, years)	-	-									
Selection strategy (e.g. inclusion/exclusion criteria)	-					-					
Providing list of included/excluded studies	-	-			-	-		-	-		-
Duplicate study selection and data extraction/control of bias		-						-	-		
Validity/quality assessment											
Using the Validity/quality assessment in conclusions	-	-				-	-	-			
Data extraction process	-	-	-	-	-			-			
Summarizing studies' characteristics		-						-	-		
Checking the combinability of the results											-
Quantitative data synthesis			-	-	-			-			
Likelihood of publication bias		-						-			
Additional analyses (e.g. sensitivity or subgroup analysis or meta-regression)		-	-	-	-			-	-		
Stating conflict of interest	-	-			-	-		-	-		-

or might help users (e.g. clinicians, practitioners, and decision-makers) to appraise the quality of MA/SR that they are referencing. Because there is no validated tool to assess specifically the methodological quality of SCED MAs and SRs, applied researchers can use one of the existing tools or combinations of multiple tools, or better yet develop and validate a new tool to conduct high-quality MAs/SRs of SCED studies' results. Most of these tools could be used to evaluate the quality of SCED MAs/SRs because they mainly focus on general aspects of the methodological quality of studies that do not depend highly on primary studies included in the review.

However, it is possible that some of the detailed criteria of the reviewed tools require being modified, omitted, or added to make it more applicable for assessing SCED MAs as was done in the study by Jamshidi et al. (in press)[19]. For instance, based on the recommendations of *What Works Clearinghouse* (WWC)[2] for combining the results of multiple SCED studies into a single summary, MAs have to meet certain thresholds: 1) a minimum of five SCED studies examining the intervention that *Meet Evidence Standards or Meet Evidence Standards with Reservations*, 2) the SCED studies must be conducted by at least three different research teams at three different geographical locations, and 3) the aggregated number of experiments across the studies must total at least 20. Such criteria can help SCED meta-analysts ensure they are following some standards while conducting their own reviews. In addition, the possible features of SCED data such as time trend or serial dependency that might lead to overestimated or underestimated intervention effect should be taken into consideration in meta-analyses. None of the reviewed tools specifically took into account these SCED-specific recommendations and it might be because the tools were not developed for assessing the quality of SCED MAs in particular. These recommendations can be considered by either the meta-analysts or users while developing new tools or applying existing tools for assessing the methodological quality of SCED MAs.

Acknowledgement

Conflicts

Authors declare that they have no conflict of interest.

References

1. Barlow DH, Nock MK, Hersen M. *Single Case Experimental Designs*. 3rd ed. Boston: Pearson/Allyn and Bacon; 2009.

2. Kratochwill TR, Hitchcock J, Horner RH, et al. Single-case design technical documentation. What Works Clearing House. http://ies. ed.gov/ncee/wwc/pdf/wwc_scd.pdf. Published 2010.

3. Onghena P. Single-case designs. In: Everittt BS, Howell DC, eds. *Encyclopedia of Statistics in Behavioral Science*. Chichester: John Wiley & Sons; 2005:1850-1854.

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

4. Beretvas SN, Chung H. An evaluation of modified R 2-change effect size indices for single-subject experimental designs. *Evid Based Commun Assess Interv*. 2008; 2(3): 120-128. doi:10.1080/17489530802446328.

5. Moeyaert M, Ugille M, Ferron JM, et al. The influence of the design matrix on treatment effect estimates in the quantitative analyses of single-subject experimental design research. *Behav Modif*. 2014; 38(5): 665-704. doi:10.1177/0145445514535243.

6. Rogers LA, Graham S. A meta-analysis of single subject design writing intervention research. *J Educ Psychol*. 2008; 100(4): 879-906. doi:10.1037/0022-0663.100.4.879.

7. Smith JD. Single-case experimental designs: A systematic review of published research and current standards. *Psychol Methods*. 2012; 17(4): 1-70. doi:10.1037/a0029312.

8. Schlosser RW, Lee DL, Wendt O. Application of the percentage of non-overlapping data (PND) in systematic reviews and meta-analyses: A systematic review of reporting characteristics. *Evidence-Based Commun Assess Interv*. 2008; 2(3): 163-187. doi:10.1080/17489530802505412.

9. Shadish WR. Statistical analyses of single-case designs: The shape of things to come. *Curr Dir Psychol Sci*. 2014; 23(2): 139-146. doi:10.1177/0963721414524773.

10. Shadish WR, Hedges LV. Pustejovsky JE. Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: A primer and applications. *J Sch Psychol*. 2014; 52(2): 123-147. doi:10.1016/j.jsp.2013.11.005.

11. Shadish WR, Rindskopf DM. Methods for evidence-based practice: Quantitative synthesis of single-subject designs. *New Dir Eval*. 2007; 113: 95-109. doi:10.1002/ev.217.

12. Petit-Bois M, Baek EK, Van den Noortgate W, et al. The consequences of modeling autocorrelation when synthesizing single-case studies using a three-level model. *Behav Res Methods*. 2016; 48(2): 803-812. doi:10.3758/s13428-015-0612-1.

13. Tincani M, De Mers M. Meta-analysis of single-case research design studies on instructional pacing. *Behav Modif*. 2016; 40(6): 799-824. doi:10.1177/0145445516643488.

14. Petticrew M, Roberts H. *Systematic Reviews in the Social Sciences: A Practical Guide*. Oxford: Blackwell Publishing Limited; 2006. doi:10.1002/9780470754887.

15. Cooper H, Hedges LV. Research synthesis as a scientific process. In: Cooper H, Hedges L V., Valentine JC, eds. *Handbook of Research Synthesis and Meta-Analysis*. 2nd ed. New York: Russell Sage Foundation; 2009:3-18. http://www.google.se/books?hl=sv&lr=&id=LUGd6B9eyc4C&pgis=1.

16. Campbell JM. Statistical comparison of four effect sizes for single-subject designs. *Behav Modif*. 2004; 28(2): 234-246. doi:10.1177/0145445503259264.

17. Van den Noortgate W, Onghena P. Hierarchical linear models for the quantitative integration of effect sizes in single-case research. *Behav Res methods instruments Comput*. 2003; 35(1): 1-10. doi:10.3758/BF03195492.

18. Van den Noortgate W, Onghena P. A multilevel meta-analysis of single-subject experimental design studies. *Evid Based Commun Assess Interv*. 2008; 2(3): 142-151. doi:10.1080/17489530802505362.

19. Jamshidi L, Heyvaert M, Declercq L, et al. Methodological quality of meta-analyses of single-case experimental studies. Res Dev Disabil. doi:10.1016/j.ridd.2017.12.016.

20. Remschmidt C, Wichmann O, Harder T. Methodological quality of systematic reviews on influenza vaccination. *Vaccine*. 2014; 32(15): 1678-1684. doi:10.1016/j.vaccine.2014.01.060.

21. Faggion CMJ, Giannakopoulos NN. Critical appraisal of systematic reviews on the effect of a history of periodontitis on dental implant loss. *J Clin Periodontol*. 2013; 40(5): 542-552. doi:10.1111/jcpe.12096.

22. Pinnock H, Parke HL, Panagioti M, et al. Systematic meta-review of supported self-management for asthma: A healthcare perspective. *BMC Med*. 2017; 15(64). doi:10.1186/s12916-017-0823-7.

23. Wells C, Kolt GS, Marshall P, Hill B, et al. Effectiveness of Pilates exercise in treating people with chronic low back pain: A systematic review of systematic reviews. *BMC Med Res Methodol*. 2013; 13(7): 1-12. doi:10.1186/1471-2288-13-7.

24. Hall AM, Lee S, Zurakowski D. Quality assessment of meta-analyses published in leading anesthesiology journals from 2005 to 2014. *Anesth Analg*. 2017; 124(6): 2063-2067. doi:10.1213/ANE.0000000000002074.

25. Rotta I, Salgado TM, Silva ML, et al. Effectiveness of clinical pharmacy services: An overview of systematic reviews (2000–2010). *Int J Clin Pharm*. 2015; 37(5): 687-697. doi:10.1007/s11096-015-0137-9.

26. Pieper D, Mathes T, Eikermann M. Can AMSTAR also be applied to systematic reviews of non-randomized studies? *BMC Res Notes*. 2014; 7(609): 1-6. doi:10.1186/1756-0500-7-609.

27. Faggion CMJ. Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor. *BMC Med Res Methodol*. 2015; 15(63). doi:10.1186/s12874-015-0062-6.

28. Shea BJ, Grimshaw JM, Wells GA, et al. Development of AMSTAR: A measurement tool to assess the methodological quality of systematic reviews. *BMC Med Res Methodol*. 2007; 7(10): 1-7. doi:10.1186/1471-2288-7-10.

29. Zeng X, Zhang Y, Kwong JSW, et al. The methodological quality assessment tools for preclinical and clinical studies , systematic review and meta-analysis , and clinical practice guideline : A systematic review. 2015; 8: 2-10. doi:10.1111/jebm.12141.

30. Pussegoda K, Turner L, Garritty C, et al. Systematic review adherence to methodological or reporting quality. *Syst Rev*. 2017; 6(131): 1-14. doi:10.1186/s13643-017-0527-2.

31. Pussegoda K, Turner L, Garritty C, et al. Identifying approaches for assessing methodological and reporting quality of systematic reviews: A descriptive study. *Syst Rev*. 2017; 6(1): 1-13. doi:10.1186/s13643-017-0507-6.

32. Sacks HS, Berrier J, Reitman D, et al. Meta-analyses of randomized controlled trials. *N Engl J Med*. 1987; 316(8): 450-455. doi:10.1056/NEJM198702193160806.

33. Oxman AD, Guyatt GH, Singer J, et al. Agreement among reviwers of review articles. *J Clin Epidemiol*. 1991; 44(1): 91-98. doi:https://doi.org/10.1016/0895-4356(91)90205-N.

34. Salmos J, Gerbi MEMM, Braz R, et al. Methodological quality of systematic reviews analyzing the use of laser therapy in restorative dentistry. *Lasers Med Sci*. 2010; 25(1): 127-136. doi:10.1007/s10103-009-0733-9.

35. Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence- based health care: Validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. *Open Dent J*. 2010; 4: 84-91. doi:10.2174/1874210601004020084.

36. *Methodology Checklist 1: Systematic Reviews and Meta-Analyses*. Scottish Intercollegiate Guidelines Network; 2012. http://www.sign.ac.uk/checklists-and-notes.html.

37. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *Br J Surg*. 2000; 87: 1448-1454. doi:10.1159/000055014.

38. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *J Chinese Integr Med*. 2009; 7(9): 889-896. doi:10.3736/jcim20090918.

39. Critical Appraisal Skills Programme. CASP Systematic Review

Jamshidi L, Declercq L, Ferron JM, Moeyaert M, Beretvas SN, Noortgate WN. Improving the Methodological Quality of Single-Case Experimental Design Meta-Analysis. J Ment Health Clin Psychol (2018) 2(4): 1-9

Journal of Mental Health & Clinical Psychology

checklist. https://casp-uk.net/casp-tools-checklists/. Published 2018.

40. *The Social Care Guidance Manual:Process and Methods.* National Institute for Health and Care Excellence; 2013. https://www.nice.org.uk/process/pmg10/chapter/introduction. Accessed April 13, 2018.

41. Churchill R, Lasserson T, Chandler J, et al. Standards for the reporting of new Cochrane Intervention Reviews. In: Higgins JP, Lasserson T, Chandler J, Tovey D, Churchill R, eds. *Methodological Expectations of Cochrane Intervention Reviews*. Cochrane: London; 2016:37-58.

42. Whiting P, Savović J, Higgins JPT, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016; 69: 225-234. doi:10.1016/j.jclinepi.2015.06.005.