

Group	Quality criteria	Sub-criteria	Is criterion applied in Meta-analysis (to what extent)	Score	Description	Reference	
Literature search and inclusion / exclusion criteria	1. Literature search	Published literature extracted from	> 4 databases	3			
			between 2 and 4 databases	2	Several databases should be used for extracting published literature to reduce the risk of selection bias	Côté et al. (2013, p. 40)	
			1 database	1			
			not reported	0			
		Grey literature (unpublished reports and experiments, project reports etc.) included	yes	1	Grey literature maximizes comprehensiveness and reduces risk of bias. Whether conducting a grey literature search is necessary or not is dependent on the meta-analysis itself and needs to be assessed by the authors	Borenstein et al. (2009, p. 280)	
	2. Authors checked the reference lists of other existing meta-analyses and reviews for available literature	Keywords/search string reported	yes	1	The search string(s) used to retrieve literature from different databases should be stated	Côté et al. (2013, p. 43)	
			no	0			
	3. Inclusion and exclusion criteria reported		yes	1	Ensures the inclusion of more relevant articles, as occasionally, keyword searching in databases does not provide results for all available literature	Borenstein et al. (2009, p. 278)	
			no	0			
	4. Control (C) and treatment (T) described		yes	1	Inclusion of studies on the same research topic and avoiding mixing "apples and oranges". The controls of included studies should be comparable to allow the correct computation of the treatment and moderator effect	Stewart et al. (2013, p. 28)	
		no	0				
5. Moderators and their ranges or groups described		yes	1	Defining moderators is essential to evaluate the source of variation across studies. Their ranges or groups are important to indicate the limits within which moderators were studied	Stewart et al. (2013, p. 32)		
		no	0				
Cutoff criteria	6. Effect size	In(R) (log response ratio)		2	Easily interpretable by back-transformation of $\ln(R)$ to a percentage change from the control $\ln(R) = \ln(\bar{X}_1) - \ln(\bar{X}_2)$	Rosenberg et al. (2013, p. 63f)	
		Raw mean difference (D)		1	Not recommended for meta-analyses having a range of control levels/scales. Example: when SOC stocks are studied, initial values can have a wide range (10-100 t ha ⁻¹) $D = \bar{X}_1 - \bar{X}_2$	Borenstein et al. (2009, p. 21ff)	
	Standardized mean difference (e.g., Hedges' d)		1	Difficult to interpret: $ d = 0.2$ - small effects $ d = 0.5$ - moderate effects $ d = 0.8$ - large effects	Borenstein et al. (2009, p. 26); Rosenberg et al. (2013, p. 63f)		
	Non-standard metrics used or not calculated		0	Without the calculation of effect sizes, a synthesis does not qualify as a meta-analysis (all following quality criteria of group "Meta-analysis" account for 0)	Koricheva and Gurevitch (2014, p. 840)		
	7. Standard deviation extracted (or computed from statistics)		From each study	2	Standard deviations (SDs) need to be extracted (or computed from available statistics, see Acutis et al. 2022) from all studies in order to calculate the weight for each study. If SDs were extracted only from some studies and the rest was ignored or roughly estimated by, e.g., calculating the overall mean SD of the available SD, a score of "1" was assigned	Nakagawa et al. (2017, p.11)	
		From some studies	1				
		Not extracted	0				
8. Studies weighted by 1/variance		For each study	4	Weighting of studies is only correct when done by the inverse of variance. Meta-analyses that extracted SD only for some studies but weighted by 1/variance did weigh "some studies" or "partly"	Koricheva and Gurevitch (2013a, p. 9)		
		For some studies	1				
		Not weighted / reported	0	$W_i = \frac{1}{V_i}$			
9. Subgroup analysis and meta-regression		yes	1	Categorical and/or continuous moderators should be assessed by χ^2 test	Rosenberg et al. (2000, p. 111f)		
		no	0				
	10. Model used	Random-effect or mixed-effect model	1	When conducting a meta-analysis in the field of agriculture, the random-effect or mixed-effect model should be chosen, as it acknowledges between-study variation	Mengersen et al. (2013, p. 94)		
	Fixed-effect model	0					
	No model reported	0					
11. Software used for meta-analysis		Meta-analytical software (e.g., MetaWin, Metafor package, RevM, or other software (e.g., SPSS, SAS, Stata, R, etc.))	1	Used software should be stated; when using general statistical analysis software, correct model choice (weighted + random model) and implementation are necessary	Schmid et al. (2013, p. 174)		
		Spreadsheet (as MS Excel) or not reported	0				
	12. Effect sizes are independent		1-2 effect sizes per study/site extracted	1	Effect sizes should be independent. For each study or site, one effect size should be extracted. When several combinations of treatment and control were studied, only one effect size per study or site should be extracted	Gurevitch and Hedges (1999, p. 1147); Hungate et al. (2009, p. 2009f); Nakagawa et al. (2017, p. 3)	
		> 2 effect sizes per study/site extracted	0				
13. Response variables and relevant parameters for their calculation were measured		yes	1	Response variables need to be measured in an experiment, not estimated or modeled (e.g., pedotransfer functions only provide estimates for bulk density and therefore, introduce the risk of inaccurate calculation of SOC stock contents)	Xu et al. (2015, p. 1574)		
		no	0				
14. Sensitivity analysis to test robustness of meta-analysis	Outliers and effect size distribution	yes	1	Presence of outliers should be tested and can be identified via effect size distribution in weighted histograms, box plots, etc.	Rothstein et al. (2013, p. 333)		
	Testing for publication bias	yes	1	Magnitude of publication bias should be estimated by funnel plots, Egger's regression, or Fail-safe test	Borenstein et al. (2009, p. 291)		
		no	0				
15. Results presentation in figures and tables	Summarized effect size and Confidence Intervals	yes	1	Summarized effect size and confidence intervals should be presented in a table or figure	Borenstein et al. (2009, p. 6)		
	Moderator analysis (sub-group analysis and/or meta-regression)	yes	1	Moderator (covariates) analysis should be presented in the form of figures or tables	Borenstein et al. (2009, p. xxii)		
		no	0				
	Forest plot	yes	1	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	Lortie et al. (2013, p. 344f)		
		no	0				
16. Description of meta-data		yes	1	A description of meta-data (authors and year plus experimental location, treatments, etc.) in the article or appendix should be provided as a table	Gonçalves and Musen (2019, p. 2)		
		no	0				
Results and database presentation	17. Full database including most of the following criteria for each study	Article ID and/or first author plus year		-	Each study should have a distinctive number to allow easy identification	-	
		Country/location		-	First author and year of publication should be stated	-	
		Control (C)	Available and includes most of the listed criteria: Article ID and/or first author plus year must be described; either effect size or mean, SD and sample size of T/C must be described		-	Country and exact location of experiment	-
		Treatment (T)			-	Measure used as control	-
		Moderators		2	Possible moderators (ped-climatic, experimental conditions, duration of experiments, land use/crops/cropping systems)	Measure under investigation	-
		Means of C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-	
		Standard deviations of C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-	
		Sample sizes C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-	
		Effect size		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-	
		Article ID and/or first author plus year	Available and includes the listed criteria (either effect size or mean, SD and sample size of T/C must be described)	1	see above in point 17	see above in point 17	
	Means of C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-		
	Standard deviations of C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-		
	Sample sizes C/T		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-		
	Effect size		-	Forest plots enable the graphical presentation of individual effect sizes of studies and the overall effect size including confidence intervals	-		
	Not available or includes less information than score 1		0	If no database is provided, the meta-analysis is not transparent	-		
	Maximum reachable score per meta-analysis:		30				