

Systematic reviews and meta-analysis

Criterion	Green	Yellow	Red	Comments
The study is in fact identified as a systematic review or meta-analysis	“Systematic review,” “meta-analysis,” or both, are in the title of the article, and the abstract supports the design in the title	The title is ambiguous, but the abstract shows that the authors did a systematic review	The article is a narrative review or an overview, or is done by a single author	“Systematic review” and “meta-analysis” are generally recognized terms for a specific type of original research; narrative reviews are subject to biases which systematic reviews and meta-analyses methodically control for
Objectives of the systematic review or meta-analysis	Clearly stated in terms of P ICOS: P atient population (disease, age, setting), I ntervention (dose, frequency, etc), C omparator (control group interventions), O utcome (morbidity, mortality, symptoms, function), and S tudy design (randomized trials only, broader design criteria)	PICOS elements all reported, but some ambiguity in some elements (e.g., Comparator described as “standard care” or “usual care” without further description)	One or more PICOS element missing or uninterpretable	The question being addressed should be clear from the abstract; it may be narrow or broad, but the scope and potential applicability should be well defined
Characteristics of eligible studies	In addition to PICOS, study characteristics	Ambiguity exists for some of the	Eligibility of studies is unclear, and	

Criterion	Green	Yellow	Red	Comments
	defined in terms of restrictions for inclusion (e.g., minimum length of follow-up, whether co-interventions are included), and scope of reports (language, years of publication, unpublished material)	characteristics of eligible studies	scope of reports is not specified	
Information sources	Multiple information sources are clearly specified: databases (PubMed, Ovid, EMBASE, Cochrane, Web of Science), hand searches of tables of contents of relevant journals, meeting abstracts, reference lists, contacts with authors, manufacturers, trial registries)	Search limited to published material from two or more sources, without additional searching of registries or contact with authors	Search limited to a single information source (e.g., PubMed only)	While PubMed is a large and nearly comprehensive database, its yield can be influenced by how articles are indexed by the National Library of Medicine; additional sources of information can materially affect the conclusions of a systematic review or meta-analysis
Search strategy	Full electronic search strategy for at least one major database, with dates (e.g., PubMed 1970-	Databases and search terms are given, but there is some ambiguity in the strategy	Databases and search terms are too broad and vague to permit replication by	Often given in an appendix to the article or in an online supplement, the strategy should

Criterion	Green	Yellow	Red	Comments
	October 2009), limits, combinations of search terms, such that it can be replicated by the reader	(e.g., PubMed “through 2007”), and replication by the reader would be difficult	the reader	be readily accessible
Study selection	Specification of which criteria determine eligibility for inclusion (e.g., randomization to specified interventions, which outcomes were required to be reported) and for quality (e.g., allocation concealment, intention-to-treat analysis, blinding) with at least two reviewers identified by initials; inter-rater agreement and methods of resolving disagreement are specified; a flow diagram enumerates articles retrieved from search, articles excluded after screening, and articles included for meta-analysis	Two or more reviewers screen articles for inclusion, but there is some ambiguity in the criteria for inclusion or for inter-rater agreement and methods of resolving disagreement; flow diagram is lacking	Only one reviewer selects studies; criteria are vague	Quality assessment should focus on risk of bias; scoring of articles for quality is not necessary and may be misleading. There is no standard process for selecting studies, but the process used by the reviewers should be clear enough to allow the reader to determine which studies might meet the test of inclusion
Outcomes for	Meta-analysis	Meta-analysis	Meta-analysis	Exploratory

Criterion	Green	Yellow	Red	Comments
analysis	is restricted to pre-specified primary and secondary outcomes, and exploratory (hypothesis-generating) analyses in the source literature are excluded from meta-analysis	combines pre-specified primary and secondary outcomes in the source literature with exploratory analyses in the same literature, but assigns exploratory analyses a lower weight	treats exploratory analyses in source literature on an equal basis with the pre-specified primary and secondary analyses	analyses are too likely to be reported when they arise from the play of chance, and should not be included in any meta-analysis of the same outcomes; their inclusion is likely to bias the meta-analysis
Summary measures for meta-analysis with or without pooled Number Needed to Treat (NNT)	Principal summary measures (relative risk, risk difference, odds ratio, difference in means, hazard ratio) are specified and appropriate to the outcome measure; if numbers needed to treat (NNT) are reported, there is a fixed event rate in the control groups for the studies being combined	Risk ratios or odds ratios are reported, and NNT is not reported if there is a difference in the control group event rates across the different studies	Risk ratios or odds ratios are reported, but NNT is reported even when there is a difference in control group event rates across the different studies (the underlying baseline risks are not equal)	Relative risks and odds ratios are generally more stable for summary measures than risk differences; pooled NNT is misleading if the control group event rate (the baseline risk) is different across studies, even if the risk ratio is the same
Meta-analysis presentation	Results of meta-analysis are presented as an estimated summary effect (with confidence interval) across	Estimated summary effect with confidence interval, with an estimate of heterogeneity, and an explanation of	Summary effect measure with confidence interval, but heterogeneity measures and examinations are lacking	No hard and fast rule dictates the choice of model, but because a fixed effect model assumes a

Criterion	Green	Yellow	Red	Comments
	all included studies, displaying a forest plot with weights and confidence intervals for the included studies; a measure of heterogeneity is presented (e.g., I^2); the choice of fixed effect or random effects model is explained, and, if there is significant heterogeneity, there is an attempt to examine possible sources of heterogeneity	the choice of fixed or random effects model; however, an examination of sources of heterogeneity is lacking		single common effect size across studies, there should be a discussion of why it is appropriate for the included studies