

Table S5. General characteristics of studies evaluating the measurement properties of tools for assessing risk of reporting biases

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Armijo-Olivo 2012 ¹	Cochrane risk of bias tool for randomized trials (2008 version)	None	20 trials included in a SR exploring knowledge transfer interventions for cancer pain management.	Cancer pain	None	20	NA	Range 1987-2007	2
Armijo-Olivo 2014 ²	Cochrane risk of bias tool for randomized trials (2011 version)	Inter-rater reliability	Trials of physical therapy interventions included in meta-analyses of a continuous outcome.	Physical therapy for musculoskeletal, cardiorespiratory, neurological or gynaecological conditions	None	109	NA	Not reported	2
Bilandzic 2016 ³	ROBINS-I (Risk Of Bias In Non-randomized Studies of Interventions) tool	Inter-rater reliability	Studies included in two SRs of NRSI of the relationship between the use of TZDs and COX-2 inhibitors and major cardiovascular events.	Cardiovascular disease	None	37	NA	Range 2000-2010	2

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Downs 1998 ⁴	Downs-Black tool	None	10 randomised controlled trials and 10 non-randomised trials/prospective cohort studies randomly selected from studies identified during a SR of surgery for stress incontinence	Stress incontinence	None	20	NA	Not reported	2
Hartling 2009 ⁵	Cochrane risk of bias tool for randomized trials (2008 version)	Inter-rater reliability	A convenience sample of 163 randomized trial in child health, which were presented at the annual scientific meetings of the Society for Pediatric Research between 1992 and 1995.	Child health	None	163	NA	Not reported	2

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Hartling 2011 ⁶	Cochrane risk of bias tool for randomized trials (2008 version)	Inter-rater reliability	Trials included in a systematic review of long-acting beta agonists (LABA) combined with inhaled corticosteroids (ICS) for adults with persistent asthma.	Asthma	None	107	NA	Median 2004, IQR 2001-2006	2
Hartling 2012 ^{7,8}	Cochrane risk of bias tool for randomized trials (2011 version)	Inter-rater reliability	A sample of 154 trial was randomly selected from among 616 trials published in December 2006 that were previously examined for quality of reporting.	Varied	None	154	NA	All 2006	2
Hayden 2013 ⁹	QUIPS (Quality In Prognosis Studies) tool	Inter-rater reliability	Studies included in a systematic review of troponin-based risk stratification of patients with	Pulmonary embolism	None	31	NA	Not reported	2

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
			acute non-massive pulmonary embolism.						
Hoojimans 2014 ¹⁰	SYRCLE's RoB tool (Systematic Review Centre for Laboratory animal Experimentation)	Inter-rater reliability	1 systematic review including 32 papers (no other details provided).	Animal studies (not specified)	None	32	NA	Not reported	2
Jordan 2017 ¹¹	Cochrane risk of bias tool for randomized trials (2011 version)	Inter-rater reliability	Any study that had been included more than once in SRs present on the Cochrane Database of Systematic Reviews in the area of subfertility.	Subfertility	None	28	NA	Not reported	2
Kim 2013 ¹²	RoBANS (Risk of Bias Assessment Tool for Nonrandomized Studies)	Inter-rater reliability	39 NRSs from four systematic reviews (one by the National Evidence-based Healthcare Collaborating Agency and three Cochrane reviews).	Depression, myocardial infarction, post-partum hemorrhage, chronic non-cancer pain	None	39	NA	Not reported	2

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Kumar 2016 ¹³	GRADE	None	10 key questions that were systematically reviewed for a clinical practice guideline for the use of prophylactic vs. therapeutic platelet transfusion in patients with thrombocytopenia.	Thrombocytopenia	10	None	All 2015	NA	18
Llewellyn 2015 ¹⁴	SAQAT (Semi-Automated Quality Assessment Tool)	Inter-rater reliability	29 meta-analyses from a purposive sample of SRs of RCTs from the Database of Systematic Reviews of Effects (DARE), and a purposive sample of 15 recent Cochrane reviews in mental health.	Varied	44	None	2006-2013	NA	2
Mustafa 2013 ¹⁵	GRADE	None	4 well-conducted and well-reported Cochrane reviews,	Alcohol dependence, asthma,	16	None	2004-2012	NA	4

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Norris 2012 ¹⁶	ORBIT-I (Outcome Reporting Bias In Trials) classification system for benefit outcomes	Inter-rater reliability; Time to complete assessments	based on assessment using the AMSTAR tool. Studies included in three AHRQ-funded comparative effectiveness reviews of randomised trials with drug-drug or drug-placebo comparisons, examining benefit outcomes.	cardiopulmonary bypass Varied	None	40	NA	2005-2010	2
O'Connor 2015 ¹⁷	Downs-Black tool	None	20 studies included in an updated SR which examined the effects of an exercise intervention for chronic musculoskeletal pain.	Chronic musculoskeletal pain	None	20	NA	1997-2008	2

Study ID	Tool assessed	Properties evaluated for reporting bias item	Sampling frame	Areas of health care	No. syntheses assessed	No. studies assessed	Publication years of syntheses	Publication years of studies	No. assessors
Vale 2013 ¹⁸	Cochrane risk of bias tool for randomized trials (2011 version)	Agreement between assessments performed using published article only versus published article and data collected during the individual participant data process.	13 completed individual participant data meta-analyses of treatments for cancer. Trials had to be published either in full or as an abstract, and a copy of the trial protocol or forms detailing trial design completed by trialists (or both) had to be available.	Cancer pain	None	95	NA	Not reported	2

NA = Not applicable; SR = systematic review

References

1. Armijo-Olivo S, Stiles CR, Hagen NA, et al. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract* 2012;18(1):12-8.
2. Armijo-Olivo S, Ospina M, da Costa BR, et al. Poor reliability between Cochrane reviewers and blinded external reviewers when applying the Cochrane risk of bias tool in physical therapy trials. *PLoS One* 2014;9(5):e96920.
3. Bilandzic A, Fitzpatrick T, Rosella L, et al. Risk of Bias in Systematic Reviews of Non-Randomized Studies of Adverse Cardiovascular Effects of Thiazolidinediones and Cyclooxygenase-2 Inhibitors: Application of a New Cochrane Risk of Bias Tool. *PLoS Med* 2016;13(4):e1001987.
4. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health* 1998;52(6):377-84.
5. Hartling L, Ospina M, Liang Y, et al. Risk of bias versus quality assessment of randomised controlled trials: cross sectional study. *BMJ* 2009;339:b4012.
6. Hartling L, Bond K, Vandermeer B, et al. Applying the risk of bias tool in a systematic review of combination long-acting beta-agonists and inhaled corticosteroids for persistent asthma. *PLoS One* 2011;6(2):e17242.
7. Hartling L, Hamm M, Milne A, et al. AHRQ Methods for Effective Health Care. Validity and Inter-Rater Reliability Testing of Quality Assessment Instruments. Rockville (MD): Agency for Healthcare Research and Quality (US) 2012.
8. Hartling L, Hamm MP, Milne A, et al. Testing the risk of bias tool showed low reliability between individual reviewers and across consensus assessments of reviewer pairs. *J Clin Epidemiol* 2013;66(9):973-81.

9. Hayden JA, van der Windt DA, Cartwright JL, et al. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158(4):280-6.
10. Hooijmans CR, Rovers MM, de Vries RBM, et al. SYRCLE's risk of bias tool for animal studies. *BMC Med Res Methodol* 2014;14:43.
11. Jordan VM, Lensen SF, Farquhar CM. There were large discrepancies in risk of bias tool judgments when a randomized controlled trial appeared in more than one systematic review. *J Clin Epidemiol* 2017;81:72-76.
12. Kim SY, Park JE, Lee YJ, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *J Clin Epidemiol* 2013;66(4):408-14.
13. Kumar A, Miladinovic B, Guyatt GH, et al. GRADE guidelines system is reproducible when instructions are clearly operationalized even among the guidelines panel members with limited experience with GRADE. *J Clin Epidemiol* 2016;75:115-8.
14. Llewellyn A, Whittington C, Stewart G, et al. The Use of Bayesian Networks to Assess the Quality of Evidence from Research Synthesis: 2. Inter-Rater Reliability and Comparison with Standard GRADE Assessment. *PLoS One* 2015;10(12):e0123511.
15. Mustafa RA, Santesso N, Brozek J, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. *J Clin Epidemiol* 2013;66(7):736-42; quiz 42.e1-5.
16. Norris SL, Holmer HK, Ogden LA, et al. AHRQ Methods for Effective Health Care. Selective Outcome Reporting as a Source of Bias in Reviews of Comparative Effectiveness. Rockville (MD): Agency for Healthcare Research and Quality (US) 2012.
17. O'Connor SR, Tully MA, Ryan B, et al. Failure of a numerical quality assessment scale to identify potential risk of bias in a systematic review: a comparison study. *BMC Res Notes* 2015;8:224.

18. Vale CL, Tierney JF, Burdett S. Can trial quality be reliably assessed from published reports of cancer trials: evaluation of risk of bias assessments in systematic reviews. *BMJ* 2013;346:f1798.