

Appendix: Quality scores for studies included in systematic review and meta-analysis

	1	2	3	4	5	6	7	8
	Is the hypothesis/aim/objective of the study clearly described?	Are the main outcomes to be measured clearly described in the introduction or methods section?	Are the characteristics of the patients included in the study clearly described?	Are the distributions of principal confounders in each group of subjects to be compared clearly described?	Are the main findings of the study clearly described?	Does the study provide estimates of the random variability in the data for the main outcomes?	Have the characteristics of patients lost to follow-up been described?	Have actual probability values been reported (e.g., 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001?
	Yes = 1 No = 0	Yes = 1 No = 0	Yes = 1 No = 0	Yes = 2 Part = 1 No = 0	Yes = 1 No = 0	Yes = 1 No = 0	Yes = 1 No = 0	Yes = 1 No = 0
Reference (year)								
Breuer <i>et al.</i> (2014)	1	1	1	1	1	1	NA	1
Kang <i>et al.</i> (2009)	1	0	1	1	1	1	NA	1
Kwon <i>et al.</i> (2016)	1	1	1	1	1	1	NA	1
Langan <i>et al.</i> (2014)	1	1	0	2	1	1	NA	1
Lin <i>et al.</i> (2010)	1	1	1	1	1	1	NA	1
Minassian <i>et al.</i> (2016)	1	1	1	1	1	1	NA	0
Schink <i>et al.</i> (2016)	1	1	1	1	1	1	NA	NA
Sreenivasan <i>et al.</i> (2013)	1	1	1	0	1	1	NA	1
Sundström <i>et al.</i> (2015)	1	0	1	0	1	1	NA	1
Wang <i>et al.</i> (2014)	1	1	1	1	1	1	NA	1
Wu <i>et al.</i> (2015)	1	1	1	1	1	1	NA	1
Yawn <i>et al.</i> (2016)	1	1	1	1	1	1	NA	1

	9	10	11	12	13	14	15	16
	Were the subjects asked to participate in the study representative of the entire population from which they were recruited?	Were those subjects who were prepared to participate representative of the entire population from which they were recruited?	If any of the results of the study were based on "data dredging", was this made clear?	In trials and cohort studies, do the analyses adjust for different lengths of follow-up of patients?	Were the statistical tests used to assess the main outcomes appropriate?	Were the main outcome measures used accurate (valid and reliable)?	Was there adequate adjustment for confounding in the analyses from which the main findings were drawn?	Were losses of patients to follow-up taken into account?
	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0	Yes = 1 No = 0 Unable to determine = 0
Reference (year)								
Breuer <i>et al.</i> (2014)	1	1	1	1	0	1	1	1
Kang <i>et al.</i> (2009)	1	1	1	1	1	1	1	1
Kwon <i>et al.</i> (2016)	1	1	0	1	1	1	0	1
Langan <i>et al.</i> (2014)	1	1	1	1	1	1	1	1
Lin <i>et al.</i> (2010)	1	1	1	1	1	1	1	1
Minassian <i>et al.</i> (2016)	1	1	1	1	1	1	1	1
Schink <i>et al.</i> (2016)	1	1	1	1	1	1	1	1
Sreenivasan <i>et al.</i> (2013)	1	1	1	1	1	1	0	1
Sundström <i>et al.</i> (2015)	1	1	1	1	1	1	0	1
Wang <i>et al.</i> (2014)	1	1	1	1	1	1	0	1
Wu <i>et al.</i> (2015)	1	1	1	1	1	1	0	1
Yawn <i>et al.</i> (2016)	1	1	1	1	0	1	1	1

Reference (year)	Yes = 1 No = 0
Breuer <i>et al.</i> (2014)	1
Kang <i>et al.</i> (2009)	1
Kwon <i>et al.</i> (2016)	1
Langan <i>et al.</i> (2014)	1
Lin <i>et al.</i> (2010)	1
Minassian <i>et al.</i> (2016)	1
Schink <i>et al.</i> (2016)	1
Sreenivasan <i>et al.</i> (2013)	1
Sundström <i>et al.</i> (2015)	1
Wang <i>et al.</i> (2014)	1
Wu <i>et al.</i> (2015)	1
Yawn <i>et al.</i> (2016)	1