

1 The impact of blinding on trial results: A systematic review and meta- 2 analysis

3

4 Tyler Pitre

5 *Department of Medicine, McMaster University, Hamilton, ON*

6 Sarah Kirsh

7 *Department of Anesthesia, McMaster University, Hamilton, ON*

8 Tanvir Jassal

9 *Department of Anesthesia, McMaster University, Hamilton, ON*

10 Mason Anderson

11 *Faculty of Health Sciences, McMaster University, Hamilton, ON*

12 Adelia Padoan

13 *Faculty of Health Sciences, McMaster University, Hamilton, ON*

14 Alexander Xiang

15 *Faculty of Health Sciences, McMaster University, Hamilton, ON*

16 Jasmine Mah

17 *Department of Medicine, Dalhousie University, Halifax, NS*

18 Dena Zeraatkar*

19 *Department of Anesthesia; Department of Health Research Methods, Evidence, and Impact, McMaster University,*
20 *Hamilton, ON*

21

22

23 *Corresponding author:

24 **Conflicts of interest:** None.

25 **Funding:** None.

26 **Data:** Data available at <https://osf.io/qb23h/>.

27 **Acknowledgements:** None.

28 **Authors' Contributions:** DZ conceived the study. JM performed the search. SK, TJ, MA, AP, AX screened
29 studies and collected data. SK, TSP, and DZ wrote the manuscript. All authors reviewed and approved
30 the final version of the manuscript.

31 **Word count:** 3,584.

33 **Abstract**

34 **Background:** Blinding—the concealment of the arm to which participants have been randomized—is an
35 important consideration for assessing risk of bias of randomized trials. A growing body of evidence has,
36 however, yielded inconsistent results on whether trials without blinding produce biased findings.

37 **Objective:** To conduct a systematic review and meta-analysis of the evidence addressing whether trials
38 with and without blinding produce different results.

39 **Methods:** We searched MEDLINE, EMBASE, Cochrane Reviews, JBI EBP, and Web of Science, from
40 inception to May 2022, for studies comparing the results of trials with and without blinding. Pairs of
41 reviewers, working independently and in duplicate, reviewed search results for eligible studies and
42 extracted data.

43 We pooled the results of studies comparing trials with and without blinding of patients, healthcare
44 providers/investigators, and outcome assessors/adjudicators using frequentist random-effects meta-
45 analyses. We coded study results such that a ratio of odds ratio (ROR) < 1 and difference in standardized
46 mean difference (dSMD) < 0 indicate that trials without blinding overestimate treatment effects.

47 **Results:** We identified 47 eligible studies. For dichotomous outcomes, we found low certainty evidence
48 that trials without blinding of patients and healthcare providers, outcome assessors/adjudicators, and
49 patients may slightly overestimate treatment effects. For continuous outcomes, we found low certainty
50 evidence that trials without blinding of outcome assessors/adjudicators and patients may slightly
51 overestimate treatment effects.

52 **Conclusion:** Our systematic review and meta-analysis suggests that blinding may influence trial results in
53 select situations—albeit the findings are of low certainty and the magnitude of effect is modest. In the
54 absence of high certainty evidence suggesting that trials with and without blinding produce similar
55 results, investigators should be cautious about interpreting the results of trials without blinding.

56

57 **Background**

58 Randomized trials represent the optimal design for assessing the effectiveness of interventions.
59 Randomization ensures—more or less—that important known and unknown prognostic factors are
60 balanced between arms such that any observed differences between trial arms can be attributed to the
61 intervention under study. For decades, blinding—the concealment of the arm to which participants have
62 been randomized—has been an important consideration for assessing risk of bias of randomized trials
63 (1, 2).

64 Blinding of participants and healthcare providers reduces opportunity for differences in care (co-
65 interventions)—referred to as performance bias—whereas blinding of outcome assessors/adjudicators
66 reduces differences in the measurement and adjudication of outcomes between trial arms—called
67 ascertainment bias (3, 4). Preconceived notions about the efficacy and safety of interventions, even
68 subconsciously, may theoretically impact decisions about co-interventions and outcome expectancy (5,
69 6).

70 Though there are clear advantages to blinding, blinding requires significant resources, increases the
71 operational complexity of trials, and may not always be feasible. Funding and research organizations
72 have invested in several initiatives to streamline and simplify clinical trials, and foregoing blinding is a
73 logical step in this process (7-9). It is unclear, however, the extent to which open-label trials may
74 produce biased findings.

75 Meta-epidemiologic studies (studies that analyze the results of previous studies to address how
76 methodological characteristics of the studies influence their results) have yielded inconsistent findings
77 (10-15). In this systematic review and meta-analysis, we summarize evidence addressing the effects of
78 blinding on trial results.

79 **Methods**

80 We registered a protocol on Open Science Framework (<https://osf.io/qenky>) on June 9, 2022. We report
81 our systematic review using the PRISMA reporting guidelines and guidelines for reporting meta-
82 epidemiologic methods studies (16, 17).

83 ***Search strategy***

84 In consultation with an experienced research librarian, we searched MEDLINE, EMBASE, the Cochrane
85 Database of Systematic Reviews, JBI EBP, and Web of Science, from inception to May 2022, for studies
86 comparing the results of trials with and without blinding. We supplemented our search by reviewing the

87 reference lists of eligible studies. We did not restrict the search by date or language of publication.
88 Supplement 1 presents our search strategy.

89 **Screening and study eligibility**

90 Following training and calibration exercises to ensure sufficient agreement, pairs of reviewers, working
91 independently and in duplicate, reviewed titles and abstracts of search records and subsequently the
92 full-texts of records deemed potentially eligible at the title and abstract screening stage. Reviewers
93 resolved discrepancies by discussion, or, when necessary, by adjudication by a third party.

94 We included meta-epidemiologic studies that compared the results of trials in which either patients,
95 healthcare providers, outcome assessors/adjudicators, analysts, or manuscript writers were blinded to
96 trials in which either of these parties were not blinded. We included studies that investigated the effects
97 of blinding for trials addressing the effectiveness or safety of a treatment or management strategy and
98 excluded studies that investigated the effects of blinding in diagnostic studies. We did not restrict
99 eligibility based on the population, intervention, or outcomes or by language or date of publication.

100 We excluded single systematic reviews and meta-analyses that presented subgroup or sensitivity
101 analyses addressing blinding because single reviews are unlikely to provide precise estimates and
102 because such estimates have limited generalizability. We excluded conference abstracts and studies
103 addressing the effects of blinding in animal studies or *in vitro* studies.

104 **Data extraction and risk of bias assessments**

105 Pairs of reviewers, working independently and in duplicate, collected data and assessed the risk of bias
106 of eligible studies.

107 We collected data on study characteristics (clinical condition, number of reviews/meta-analyses and
108 trials included, type of outcome investigated), blinding characteristics (blinding status of participants,
109 healthcare providers/investigators, outcome assessors/adjudicators, or other parties), and the results
110 for studies with and without blinding from the least and most adjusted estimate. For studies in which
111 multiple adjusted estimates were presented, based on previous evidence that the effects of blinding
112 may be confounded by allocation concealment, we preferentially extracted the estimate that was
113 adjusted for allocation concealment (10, 18).

114 When reported, we extracted results separately for objective and subjective outcomes. We considered
115 outcomes objective if they were established laboratory measures and other outcomes not subject to

116 interpretation (e.g., all-cause mortality, hospitalization) and subjective if outcomes were patient-
117 reported or subject to interpretation from healthcare providers or investigators (e.g., adverse events).

118 To our knowledge, there are no risk of bias tools for meta-epidemiologic studies. We assessed the risk of
119 bias of eligible studies using ad-hoc criteria: selection bias (association between trial characteristics and
120 outcomes in trials selected for analysis), confounding (distortion of the effect of blinding on study results
121 due to the correlation of the study characteristic of interest with other study characteristics that
122 influence the results), bias in classification of trial characteristics (misclassification of blinding status),
123 bias in measurement of outcomes (errors in extraction or reporting of results that may be related to
124 blinding), and selective reporting bias (selective reporting of study results based on the magnitude,
125 statistical significance, or direction of findings). To make judgements about confounding bias, we
126 considered whether the study adjusted for allocation concealment, which has been shown to influence
127 trial results (10, 18), and to make judgements about bias in classification of trial characteristics and
128 measurement of outcomes, we considered the criteria used by the study to ascertain blinding status and
129 whether trial characteristics and outcome data were collected independently and in duplicate to reduce
130 the potential for errors (15, 19).

131 We rated risk of bias criteria as either low, probably low, probably high, or high risk of bias and
132 considered results at low risk of bias overall if all domains were rated as either low or probably low risk
133 of bias.

134 Reviewers resolved discrepancies by discussion or, when necessary, by third-party adjudication.

135 ***Data synthesis and analysis***

136 We report categorical trial characteristics as proportions and percentages and continuous characteristics
137 as medians and interquartile ranges (IQRs). We synthesize the results of studies that compared trials
138 with and without blinding of patients and healthcare providers/investigators (corresponding to 'double-
139 blind') (20), outcome assessors/adjudicators, patients, healthcare providers/investigators, statisticians,
140 and manuscript writers.

141 We performed a frequentist random-effects meta-analysis with the restricted maximum likelihood
142 (REML) heterogeneity estimator to pool the results of studies. For studies that addressed dichotomous
143 outcomes, we present results as ratios of odds ratios (ROR) and for studies that addressed continuous
144 outcomes, we present differences in standardized mean differences (dSMD) with the associated

145 confidence intervals. We recoded the data such that $ROR < 1$ and $dSMD < 0$ correspond to an
146 exaggerated effect in trials without blinding compared to trials with blinding.

147 Where possible, we meta-analyzed the least and most-adjusted results from primary studies. For studies
148 that converted dichotomous outcome measures to continuous outcome measures, or vice versa, we
149 combined them in analyses with the measure to which they were converted. When studies reported
150 multiple outcomes, we synthesized results for the outcome reported by the greatest number of trials.

151 We summarize heterogeneity using the I^2 statistic. We considered heterogeneity of 0%–40% as
152 potentially unimportant, 30%–60% as moderate, 50%–90% as substantial and 75%–100% as critical (21).
153 For analyses with 10 or more studies, we assessed for publication bias using funnel plots and the Egger’s
154 test (22).

155 We performed all analyses in R (version 4.03, R Foundation for Statistical Computing), using the *meta*
156 and *metafor* packages. Open Science Framework (<https://osf.io/qb23h/>) presents the data and code to
157 generate the results in this study.

158 Where meta-analysis is not possible, we report the range and distribution of results from primary
159 studies (23).

160 ***Assessment of the certainty of evidence and interpretation of study results***

161 We use a modified GRADE approach to assess the certainty of evidence (24). While GRADE has not been
162 developed to assess the certainty of meta-epidemiologic evidence, we anticipate that GRADE criteria are
163 still relevant to methodologic questions and meta-epidemiologic studies.

164 The GRADE approach rates the evidence as either high, moderate, low, or very low certainty based on
165 considerations of risk of bias (study limitations), inconsistency (heterogeneity in study results),
166 indirectness (differences between the questions addressed in studies and the question of interest),
167 publication bias (the tendency for studies with statistically significant results or positive results to be
168 published, published faster, or published in journals with higher visibility), and imprecision (random
169 error).

170 High certainty evidence indicates situations in which we are confident that the estimated effect
171 represents the true effect, and low or very low certainty evidence indicates situations in which the
172 estimated effect may be substantially different from the true effect.

173 We considered an ROR of 1.10 or 0.90 and a dSMD of 0.1 to indicate minimally important effects of
174 blinding on trial results, which was informed by our experiences in differences in effect estimates that
175 may influence decision-making (25, 26). Such an effect corresponds to a 10% difference in the results of
176 blinded and unblinded trials for dichotomous outcomes and a quarter standard deviation difference for
177 continuous outcomes.

178 We report our results using GRADE guidance (27), which involves describing the effect of an intervention
179 based on the certainty of evidence (i.e., high certainty evidence the exposure is effective, moderate
180 certainty evidence the exposure is probably effective, low certainty evidence the exposure may be
181 effective and very low certainty evidence the effect of the exposure is unclear).

182 **Results**

183 ***Search results***

184 Our search yielded a total of 6,623 unique records, of which 47 were eligible to be included (18, 28-73).

185 Figure 1 presents the results of the search and screening.

186 ***Study characteristics***

187 Table 1 and Supplement 2 present study characteristics. Typical studies sampled meta-analyses and
188 compared the results of trials with and without blinding using either frequentist meta-regression or
189 Bayesian hierarchical models (74, 75). Studies that used frequentist meta-regressions compared the
190 results of trials with and without blinding within each meta-analysis using meta-regression and then
191 pooled the results of coefficients corresponding to blinding across meta-analyses.

192 A minority of studies reported on studies that made within trial comparisons (37, 38, 50-52). Such
193 studies included trials in which patients were randomized to blinded or open-label substudies or studies
194 that performed both blinded and open-label assessment of outcomes.

195 Studies included a median of 26 meta-analyses and 150 trials.

196 ***Risk of bias assessments***

197 Figure 2 presents the risk of bias of studies that were included in the analysis investigating the effects of
198 blinding of patients and healthcare providers/investigators and Supplement 3 presents risk of bias
199 judgements for all studies.

200 All studies were rated at high risk of bias, primarily due to confounding and selective reporting bias.

201 Studies seldom adjusted for allocation concealment and registered or published their protocols.

202 ***Blinding of patients and healthcare providers/investigators***

203 Twenty-four studies addressed the effects of blinding of patients and healthcare providers/investigators
204 on trial results (18, 28, 29, 32, 36, 39, 40, 43-46, 49, 50, 54-56, 58-60, 64-66, 69). Table 2 presents the
205 results of our analyses.

206 For dichotomous outcomes, we meta-analyzed seven studies, including 29546 meta-analyses and 3,732
207 trials, and found evidence that trials without blinding of patients and healthcare providers may
208 overestimate dichotomous outcomes (low certainty due to risk of bias and imprecision) (Supplement 4).
209 When we restricted our analysis to studies that reported adjusted effect estimates, we found a smaller
210 difference between the results of trials with and without blinding.

211 For continuous outcomes, we meta-analyzed four studies, including 45 meta-analyses and 696 trials,
212 and found that trials without blinding may not produce different results compared to trials with blinding
213 (low certainty due to risk of bias and imprecision).

214 ***Blinding of outcome assessors***

215 Twenty-three studies addressed the effects of blinding of outcome assessors/adjudicators (30, 33-35,
216 37-39, 41, 42, 44, 51-53, 55, 57, 58, 61-63, 71-73).

217 For dichotomous outcomes, we meta-analyzed 12 studies, including 119 meta-analyses and 1,579 trials,
218 and for continuous outcomes, we meta-analyzed 7 studies, including 78 meta-analyses and 1,426 trials,
219 and found evidence that trials without blinding of outcome assessors may overestimate results (low
220 certainty due to risk of bias and imprecision). We found similar results when we restricted our analyses
221 to subjective outcomes. For dichotomous outcomes, the funnel plot showed some evidence of
222 asymmetry but the Egger test for funnel plot asymmetry was not statistically significant ($p=0.38$).

223 ***Blinding of patients***

224 Twelve studies addressed the effects of blinding of patients (31, 33-35, 41, 44, 47, 48, 55, 62, 71, 73).

225 For dichotomous outcomes, we meta-analyzed four studies, including 78 meta-analyses and 1,106 trials
226 and for continuous outcomes, we meta-analyzed seven studies, including 50 meta-analyses and 1,114
227 trials. We found evidence that trials without blinding of patients may overestimate treatment effects
228 (low certainty due to risk of bias and imprecision). The results were similar when we restricted the
229 analysis to subjective outcomes.

230 ***Blinding of healthcare providers/investigators***

231 Nine studies addressed the effects of blinding of healthcare providers (33-35, 41, 44, 55, 63, 71, 73).

232 We meta-analyzed four studies reporting on dichotomous outcomes, including 54 meta-analyses and
233 902 trials, and four studies reporting on continuous outcomes, including 18 meta-analyses and 859
234 trials. For both dichotomous and continuous outcomes, we found that trials without blinding of
235 healthcare providers may not overestimate treatment effects (low certainty due to risk of bias and
236 imprecision).

237 ***Statisticians and manuscript writers***

238 Two studies addressed the effects of blinding of statisticians, one of which did not find sufficient data,
239 and none of the studies addressed the effects of blinding of manuscript writers (34, 55).

240 The one study with data addressing the effects of blinding of statisticians found that blinding of
241 statisticians may have no effect on trial results but that the results are very uncertain (very low certainty
242 due to risk of bias, imprecision, and indirectness).

243 **Discussion**

244 ***Main findings***

245 Our systematic review presents evidence that trials without blinding may overestimate treatment
246 effects—albeit the findings are of low certainty and the magnitude of effect is modest.

247 We found inconsistencies in the effects of blinding. While we found evidence, for example, that trials
248 without blinding of patients and healthcare providers may overestimate dichotomous outcomes, we did
249 not find evidence for such an effect for continuous outcomes. Similarly, we also found inconsistencies
250 across studies. The MetaBLIND study, for example, did not find evidence that trials with and without
251 blinding produce different results—contrary to several other studies of similar magnitude included in
252 our review (40, 41, 59, 69). Further, studies that compared the effects of blinding within rather than
253 across trials showed blinding to have larger effects (38, 50-52).

254 Theoretical considerations about blinding may explain some of the observed inconsistencies. For
255 example, we anticipate several factors that may vary based on the question being investigated and that
256 may influence the effect of blinding: investigators' and patients' preconceived notions about the
257 effectiveness of the intervention; whether there are co-interventions that may influence the outcome of
258 interest; the magnitude of effect of the co-interventions on the outcome of interest; the potential for
259 imbalances in the administration of co-interventions; and the degree of interpretation involved in
260 assessment of outcomes (18, 76).

261 An open-label trial addressing a condition for which there are no effective co-interventions or a trial in
262 which healthcare providers and patients either do not have access to co-interventions or in which the
263 administration of co-interventions is strictly regimented across arms is unlikely to be at risk of
264 performance bias. Similarly, an open-label trial addressing outcomes that are objective and not subject
265 to interpretation is unlikely to be at risk of ascertainment bias.

266 In the absence of high certainty evidence suggesting that trials with and without blinding produce
267 similar results, investigators should still be cautious about open-label trials.

268 ***Strengths and limitations***

269 The strengths of our systematic review include *a priori* methods, a systematic search for eligible studies,
270 duplicate screening and data extraction, and assessment of the role of blinding across a range of clinical
271 areas. Our review improves on previous reviews by performing an assessment of risk of bias and the
272 certainty of evidence and including several recently published studies not included in previous reviews
273 (10, 15, 18, 19, 33, 40, 44).

274 Our results are limited by the risk of bias of primary studies. Nearly all studies were rated at high risk of
275 bias, primarily due to potential for confounding and selective reporting bias. There is also potential for
276 misclassification of blinding status. A substantial proportion of trials, for example, did not explicitly
277 report the blinding status for all relevant parties (77). When trials did not explicitly report blinding
278 status, studies either used validated guidance to ascertain blinding status or grouped unblinded studies
279 with studies with unclear blinding (78). Even when blinding status was reported, the reporting may not
280 have been accurate (79). Several studies used the risk of bias judgements from meta-analyses as a
281 surrogate for blinding status, which may also contain errors or may reflect reviewers' judgements
282 regarding the potential impact of blinding or lack thereof on results rather than whether the trial was
283 blinded. This potential for misclassification means that the role of blinding may be more important than
284 our estimates suggest.

285 Our results are further limited by the potential inconsistency in the direction of the effect of blinding.
286 While we anticipate that trials without blinding overestimate the beneficial effects of treatments, it is
287 possible that the direction of bias caused by lack of blinding may not be consistent. That is, in certain
288 circumstances open-label trials may overestimate effects and in other circumstances they may
289 underestimate effects.

290 All of our analyses produced results that were imprecise. Further, our estimated effects for blinding
291 were generally small and evidence users may consider such effects too small to influence decision-
292 making, though most medical treatments show modest effects and hence even small biases may impact
293 decision-making (80, 81).

294 Trials addressing questions that are sensitive to performance bias and ascertainment bias may be more
295 likely to be blinded, thus reducing observed differences between trials with and without blinding.

296 There may be overlap between trials and meta-analyses included in meta-epidemiologic studies. An
297 inspection of the search strategies and clinical areas addressed by the included studies, however,
298 suggests that this overlap is likely minimal. To reduce the potential for double-counting, when meta-
299 epidemiologic studies reported more than one estimate that would be eligible to be included in each
300 analysis, we only included the estimate from the analysis that included the greatest number of trials.

301 While we tried to pool studies addressing the effects of blinding different individuals together (e.g.,
302 patients, healthcare providers), the ways in which types of blinding were defined and operationalized
303 across studies may have been slightly different.

304 It is possible that blinding may be compromised during a trial (e.g., due to obvious adverse events,
305 poorly matched placebos, or large therapeutic effects), which may have attenuated the difference
306 between trials with and without blinding.

307 ***Implications***

308 The results of our systematic review have implications for evidence users who are appraising and
309 applying evidence from open-label trials, for clinical trialists considering the potential benefits and
310 burdens of implementing blinding in their clinical trials, and for global efforts to streamline and simplify
311 clinical trial methods (7-9).

312 Results from the MetaBLIND study encouraged a critical evaluation of the importance of blinding in
313 clinical trials (41). The study, for example, led to a revision of the Cochrane-endorsed risk of bias tool
314 (RoB 2.0) to include a more nuanced approach for assessing bias due to lack of blinding (76). Results
315 from MetaBLIND, however, were imprecise and were consistent with lack of blinding resulting in both
316 an over- and underestimation of treatment effects. Further, there is no evidence whether the new
317 approach to risk of bias assessments, proposed by the RoB 2.0 tool, that places less emphasis on
318 blinding is better correlated with trial results.

319 The imprecision in results across meta-epidemiologic studies may be driven by heterogeneity in the
320 effects of blinding across trials. That is, in some situations trials without blinding may overestimate
321 treatment effects. Our findings suggest that evidence users should be cognizant of the ways in which
322 lack of blinding may bias trial results, that clinical trialists should be judicious in their decisions to forego
323 blinding, and that evidence users should be cognizant of the ways in which lack of blinding may bias trial
324 results.

325 Evidence users should be mindful that the inability to identify evidence of when and how blinding
326 affects trial results does not mean that open-label designs are unbiased. In fact, our review shows
327 evidence that trials without blinding may overestimate treatment effects.

328 The subject of blinding invites comparisons to judgements about the quality of evidence from non-
329 randomized studies. Despite evidence that randomized trials and non-randomized studies produce
330 consistent results in many situations (82-84), we typically consider non-randomized studies to produce
331 lower certainty evidence compared to randomized trials because evidence users are unable to
332 confidently distinguish between situations in which non-randomized studies produce trustworthy or
333 untrustworthy results. Uncertainty about the effects of blinding may analogously demand always or
334 almost always rating open-label studies at high risk of bias.

335 We show that we need higher certainty evidence to determine when and how blinding affects trial
336 results. Initiatives such as Studies Within a Trial (SWATs)—research studies embedded within a host trial
337 to assess alternative ways of delivering or organizing a particular trial process—can help address this
338 need (85). While incorporating SWATs within trials to study how blinding affects results may increase
339 the costs of trials, such initiatives may prove cost-effective in the long term by simplifying and
340 streamlining future trials without compromising the trustworthiness of results.

341 **Conclusion**

342 Our systematic review and meta-analysis presents evidence that trials without blinding may
343 overestimate treatment effects—albeit the findings are of low certainty and the magnitude of effect is
344 modest. In the absence of high certainty evidence suggesting that trials with and without blinding
345 produce similar results, investigators should be cautious about interpreting the results of trials without
346 blinding.

347 **Tables**

Table 1: Table of study characteristics

Type of comparison	
Between trials	41 (87.2%)
Within trials	6 (12.8%)
Type of blinding	
Patient and healthcare providers/investigators	19 (40.4%)
Outcome assessors/adjudicators	22 (46.8%)
Patients	12 (25.5%)
Healthcare providers	9 (19.1%)
Median [IQR] number of systematic reviews/meta-analyses	26 [15, 70]
Median [IQR] number of trials	166 [35, 385.5]
Analytic model	
Bayesian hierarchical	4 (8.5%)
Frequentist meta-regression/multilevel regression	31 (66%)
Meta-analysis of trials with and without blinding	3 (6.4%)
Type of outcome	
Dichotomous	33 (70.2%)
Continuous	22 (46.8%)
Funding	
Government	21 (44.7%)
Institutional	4 (8.5%)
Industry	1 (2.1%)
Not-for-profit	2 (4.3%)
None	12 (25.5%)
Not reported	9 (19.1%)

348

Table 2: Results

Dichotomous outcomes				
Type of blinding	Studies	MAs (Trials)	ROR (95% CI)	Certainty
Patients and healthcare providers	7	296 (3732)	0.90 (0.77 to 1.05)	☐ ☐ ☐ ☐ Low (due to risk of bias, imprecision)
Patients and healthcare providers (adjusted)	5	252 (3059)	0.93 (0.87 to 0.99)	☐ ☐ ☐ ☐ Low (due to risk of bias, imprecision)
Outcome assessors/adjudicators	12	119 (1579)	0.82 (0.71 to 0.96)	☐ ☐ ☐ ☐ Low (due to risk of bias, imprecision)
Outcome assessors/adjudicators (adjusted)	1	46 (397)	1.04 (0.88 to 1.22)	☐ ☐ ☐ ☐ Very low (due to risk of bias, imprecision x2)
Patients	4	641 (1011)	0.85 (0.76 to 0.96)	☐ ☐ ☐ ☐

				Low (due to risk of bias, imprecision) ⊕⊕⊕⊕
Patients (adjusted)	1	18 (132)	0.91 (0.61 to 1.35)	Very low (due to risk of bias, imprecision x2) ⊕⊕⊕⊕
Healthcare providers/investigators	4	54 (902)	0.92 (0.79 to 1.07)	Low (due to risk of bias, imprecision) ⊕⊕⊕⊕
Healthcare providers/investigators (adjusted)	1	29 (173)	1.03 (0.68 to 1.56)	Very low (due to risk of bias, imprecision x2) ⊕⊕⊕⊕

Continuous outcomes

Type of blinding	Studies	MAs (Trials)	dSMD (95% CI)	Certainty
Patients and healthcare providers	4	45 (696)	0 (-0.1 to 0)	⊕⊕⊕⊕ Low (due to risk of bias, imprecision)
Patients and healthcare providers (adjusted)	1	35 trials	-0.2 (-0.5 to 0.1)	⊕⊕⊕⊕ Very low (due to risk of bias, imprecision x2)
Outcome assessors/adjudicators	9	90 (1511)	-0.1 (-0.2 to 0)	⊕⊕⊕⊕ Low (due to risk of bias, imprecision)
Outcome assessors/adjudicators (adjusted)	0	NA	NA	⊕⊕⊕⊕ Very low (due to risk of bias, imprecision x2)
Patients	7	50 (1017)	-0.1 (-0.4 to 0.1)	⊕⊕⊕⊕ Low (due to risk of bias, imprecision)
Patients (adjusted)	1	14 (158)	0 (-0.2 to 0.2)	⊕⊕⊕⊕ Very low (due to risk of bias, imprecision x2)
Healthcare providers/investigators	4	18 (853)	0 (-0.2 to 0.1)	⊕⊕⊕⊕ Low (due to risk of bias, imprecision)
Healthcare providers/investigators (adjusted)	2	18 (643)	0 (-0.2 to 0.3)	⊕⊕⊕⊕ Very low (due to risk of bias, imprecision x2)

dSMD: difference in standardized mean difference; ROR= ratio of odds ratio
ROR < 1 and difference dSMD < 0 indicate that trials without blinding overestimate treatment effects

349

350 References

- 351 1. Schulz KF, Grimes DA. Blinding in randomised trials: hiding who got what. *Lancet*.
352 2002;359(9307):696-700.
- 353 2. Day SJ, Altman DG. Statistics notes: blinding in clinical trials and other studies. *Bmj*.
354 2000;321(7259):504.
- 355 3. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane
356 Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928.

- 357 4. Karanickolas PJ, Farrokhyar F, Bhandari M. Practical tips for surgical research: blinding: who,
358 what, when, why, how? *Can J Surg.* 2010;53(5):345-8.
- 359 5. Nickerson RS. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General*
360 *Psychology.* 1998;2(2):175-220.
- 361 6. Rosenthal R. On the social psychology of the psychological experiment: the experimenter's
362 hypothesis as unintended determinant of experimental results. *Am Sci.* 1963;51:268-83.
- 363 7. Protas. 2022 [Available from: <https://protas.co.uk/>].
- 364 8. Corneli A, Hallinan Z, Hamre G, Perry B, Goldsack JC, Calvert SB, et al. The Clinical Trials
365 Transformation Initiative: Methodology supporting the mission. *Clin Trials.* 2018;15(1_suppl):13-8.
- 366 9. Kelly D, Spreafico A, Siu LL. Increasing operational and scientific efficiency in clinical trials. *Br J*
367 *Cancer.* 2020;123(8):1207-8.
- 368 10. Moustgaard H, Clayton GL, Jones HE, Boutron I, Jørgensen L, Laursen DRT, et al. Impact of
369 blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *Bmj.*
370 2020;368:l6802.
- 371 11. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. No firm evidence that lack of blinding
372 affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic
373 review and meta-analysis. *J Clin Epidemiol.* 2018;100:71-81.
- 374 12. Baiardo Redaelli M, Belletti A, Monti G, Lembo R, Ortalda A, Landoni G, et al. The impact of non-
375 blinding in critical care medicine trials. *J Crit Care.* 2018;48:414-7.
- 376 13. Martin GL, Trioux T, Gaudry S, Tubach F, Hajage D, Dechartres A. Association Between Lack of
377 Blinding and Mortality Results in Critical Care Randomized Controlled Trials: A Meta-Epidemiological
378 Study. *Crit Care Med.* 2021;49(10):1800-11.
- 379 14. Savovic J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association Between
380 Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-
381 Epidemiologic Study. *Am J Epidemiol.* 2018;187(5):1113-22.
- 382 15. Page MJ, Higgins JP, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical Evidence of Study
383 Design Biases in Randomized Trials: Systematic Review of Meta-Epidemiological Studies. *PLoS One.*
384 2016;11(7):e0159267.
- 385 16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020
386 statement: an updated guideline for reporting systematic reviews. *Rev Esp Cardiol (Engl Ed).*
387 2021;74(9):790-9.
- 388 17. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid*
389 *Based Med.* 2017;22(4):139-42.
- 390 18. Zeraatkar D, Pitre T, Diaz-Martinez JP, Chu D, Rochweg B, Lamontagne F, et al. Effects of
391 allocation concealment and blinding in trials addressing treatments for COVID-19: A methods study.
392 *medRxiv.* 2022:2022.08.03.22278348.
- 393 19. Dechartres A, Trinquart L, Faber T, Ravaud P. Empirical evaluation of which trial characteristics
394 are associated with treatment effect estimates. *J Clin Epidemiol.* 2016;77:24-37.
- 395 20. Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, et al. Physician
396 interpretations and textbook definitions of blinding terminology in randomized controlled trials. *Jama.*
397 2001;285(15):2000-3.
- 398 21. Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted
399 systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions.
400 *Cochrane Database Syst Rev.* 2019;10:Ed000142.
- 401 22. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple,
402 graphical test. *Bmj.* 1997;315(7109):629-34.
- 403 23. Campbell M, McKenzie JE, Sowden A, Katikireddi SV, Brennan SE, Ellis S, et al. Synthesis without
404 meta-analysis (SWiM) in systematic reviews: reporting guideline. *Bmj.* 2020;368:l6890.

- 405 24. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging
406 consensus on rating quality of evidence and strength of recommendations. *Bmj*. 2008;336(7650):924-6.
- 407 25. Hultcrantz M, Rind D, Akl EA, Treweek S, Mustafa RA, Iorio A, et al. The GRADE Working Group
408 clarifies the construct of certainty of evidence. *J Clin Epidemiol*. 2017;87:4-13.
- 409 26. Zeng L, Brignardello-Petersen R, Hultcrantz M, Mustafa RA, Murad MH, Iorio A, et al. GRADE
410 Guidance 34: update on rating imprecision using a minimally contextualized approach. *J Clin Epidemiol*.
411 2022.
- 412 27. Santesso N, Glenton C, Dahm P, Garner P, Akl EA, Alper B, et al. GRADE guidelines 26:
413 informative statements to communicate the findings of systematic reviews of interventions. *J Clin*
414 *Epidemiol*. 2020;119:126-35.
- 415 28. Ioannidis JPA, Cappelleri JC, Sacks HS, Lau J. The relationship between study design, results, and
416 reporting of randomized clinical trials of HIV infection. *Controlled Clinical Trials*. 1997;18(5):431-44.
- 417 29. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between
418 large and small randomized trials in meta-analyses. *Annals of Internal Medicine*. 2001;135(11):982-9.
- 419 30. Liu CJ, LaValley M, Latham NK. Do unblinded assessors bias muscle strength outcomes in
420 randomized controlled trials of progressive resistance strength training in older adults? *American journal*
421 *of physical medicine & rehabilitation / Association of Academic Physiatrists*. 2011;90(3):190-6.
- 422 31. Nuesch E, Reichenbach S, Trelle S, Rutjes AWS, Liewald K, Sterchi R, et al. The importance of
423 allocation concealment and patient blinding in osteoarthritis trials: A meta-epidemiologic study. *Arthritis*
424 *Care and Research*. 2009;61(12):1633-41.
- 425 32. Schulz KF, Chalmers L, Hayes RJ, Altman DG. Empirical evidence of bias: Dimensions of
426 methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the*
427 *American Medical Association*. 1995;273(5):408-12.
- 428 33. Amer MA, Herbison GP, Grainger SH, Khoo CH, Smith MD, McCall JL. A meta-epidemiological
429 study of bias in randomized clinical trials of open and laparoscopic surgery. *The British journal of*
430 *surgery*. 2021;108(5):477-83.
- 431 34. Armijo-Olivo S, Fuentes J, da Costa BR, Saltaji H, Ha C, Cummings GG. Blinding in Physical
432 Therapy Trials and Its Association with Treatment Effects: A Meta-epidemiological Study. *American*
433 *journal of physical medicine & rehabilitation*. 2017;96(1):34-44.
- 434 35. Bolvig J, Juhl CB, Boutron I, Tugwell P, Ghogomu EAT, Pardo JP, et al. Some Cochrane risk-of-bias
435 items are not important in osteoarthritis trials: a meta-epidemiological study based on Cochrane
436 reviews. *Journal of clinical epidemiology*. 2018;95:128-36.
- 437 36. Cuijpers P, Karyotaki E, Andersson G, Li J, Mergl R, Hegerl U. The effects of blinding on the
438 outcomes of psychotherapy and pharmacotherapy for adult depression: A meta-analysis. *European*
439 *psychiatry : the journal of the Association of European Psychiatrists*. 2015;30(6):685-93.
- 440 37. Dello Russo C, Cappoli N, Pilunni D, Navarra P. Local Investigators Significantly Overestimate
441 Overall Response Rates Compared to Blinded Independent Central Reviews in Phase 2 Oncology Trials.
442 *Journal of Clinical Pharmacology*. 2021;61(6):810-9.
- 443 38. Hrobjartsson A, Emanuelsson F, Thomsen ASS, Hilden J, Brorson S. Bias due to lack of patient
444 blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-
445 studies. *International Journal of Epidemiology*. 2014;43(4):1272-83.
- 446 39. Koletsi D, Spineli LM, Lempesi E, Pandis N. Risk of bias and magnitude of effect in orthodontic
447 randomized controlled trials: a meta-epidemiological review. *European journal of orthodontics*.
448 2016;38(3):308-12.
- 449 40. Martin GL, Trioux T, Gaudry S, Tubach F, Hajage D, Dechartres A. Association Between Lack of
450 Blinding and Mortality Results in Critical Care Randomized Controlled Trials: A Meta-Epidemiological
451 Study. *Critical care medicine*. 2021;49(10):1800-11.

- 452 41. Moustgaard H, Clayton GL, Jones HE, Boutron I, Jorgensen L, Laursen DLT, et al. Impact of
453 blinding on estimated treatment effects in randomised clinical trials: Meta-epidemiological study. The
454 BMJ. 2020;368:l6802.
- 455 42. Poolman RW, Struijs PAA, Krips R, Sierevelt IN, Marti RK, Farrokhyar F, et al. Reporting of
456 outcomes in orthopaedic randomized trials: does blinding of outcome assessors matter? The Journal of
457 bone and joint surgery American volume. 2007;89(3):550-8.
- 458 43. Probst P, Zschke S, Heger P, Harnoss JC, Huttner FJ, Mihaljevic AL, et al. Evidence-based
459 recommendations for blinding in surgical trials. Langenbeck's archives of surgery. 2019;404(3):273-84.
- 460 44. Saltaji H, Armijo-Olivo S, Cummings GG, Amin M, da Costa BR, Flores-Mir C. Influence of blinding
461 on treatment effect size estimate in randomized controlled trials of oral health interventions. BMC
462 medical research methodology. 2018;18(1):42.
- 463 45. Anthon CT, Granholm A, Perner A, Laake JH, Moller MH. No firm evidence that lack of blinding
464 affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic
465 review and meta-analysis. JOURNAL OF CLINICAL EPIDEMIOLOGY. 2018;100:71-81.
- 466 46. Baiardo Redaelli M, Belletti A, Monti G, Lembo R, Ortalda A, Landoni G, et al. The impact of non-
467 blinding in critical care medicine trials. Journal of Critical Care. 2018;48:414-7.
- 468 47. Braithwaite FA, Walters JL, Li LSK, Lorimer Moseley G, Williams MT, McEvoy MP. Effectiveness
469 and adequacy of blinding in the moderation of pain outcomes: Systematic review and meta-analyses of
470 dry needling trials. PeerJ. 2018;2018(7):e5318.
- 471 48. Feys F, Bekkering GE, Singh K, Devroey D. Do randomized clinical trials with inadequate blinding
472 report enhanced placebo effects for intervention groups and nocebo effects for placebo groups?
473 Systematic Reviews. 2014;3(1):14.
- 474 49. Howard JP, Shun-Shin MJ, Hartley A, Bhatt DL, Krum H, Francis DP. Quantifying the 3 Biases That
475 Lead to Unintentional Overestimation of the Blood Pressure-Lowering Effect of Renal Denervation.
476 Circulation: Cardiovascular Quality and Outcomes. 2016;9(1):14-22.
- 477 50. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Rasmussen JV, Hilden J, et al. Observer
478 bias in randomized clinical trials with time-to-event outcomes: systematic review of trials with both
479 blinded and non-blinded outcome assessors. International journal of epidemiology. 2014;43(3):937-48.
- 480 51. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias
481 in randomised clinical trials with binary outcomes: systematic review of trials with both blinded and non-
482 blinded outcome assessors. BMJ (Clinical research ed). 2012;344:e1119.
- 483 52. Hrobjartsson A, Thomsen ASS, Emanuelsson F, Tendal B, Hilden J, Boutron I, et al. Observer bias
484 in randomized clinical trials with measurement scale outcomes: a systematic review of trials with both
485 blinded and nonblinded assessors. CMAJ : Canadian Medical Association journal = journal de
486 l'Association medicale canadienne. 2013;185(4):E201-11.
- 487 53. Ndounga Diakou LA, Trinquart L, Hrobjartsson A, Barnes C, Yavchitz A, Ravaud P, et al.
488 Comparison of central adjudication of outcomes and onsite outcome assessment on treatment effect
489 estimates. Cochrane Database of Systematic Reviews. 2016;2016(3):MR000043.
- 490 54. Trone JC, Chapelle C, Ollier E, Bertolotti L, Cucherat M, Mismetti P, et al. Statistical controversies
491 in clinical research: limitations of open-label studies assessing antiangiogenic therapies with regard to
492 evaluation of vascular adverse drug events-a meta-analysis. Annals of Oncology. 2018;29(4).
- 493 55. Balk EM, Bonis PA, Moskowitz H, Schmid CH, Ioannidis JP, Wang C, et al. Correlation of quality
494 measures with estimates of treatment effect in meta-analyses of randomized controlled trials. Jama.
495 2002;287(22):2973-82.
- 496 56. Moher D, Cook DJ, Jadad AR, Tugwell P, Moher M, Jones A, et al. Assessing the quality of reports
497 of randomised trials: implications for the conduct of meta-analyses. Health Technol Assess. 1999;3(12):i-
498 iv, 1-98.

- 499 57. Jüni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-
500 analysis. *Jama*. 1999;282(11):1054-60.
- 501 58. Lega JC, Mismetti P, Fassier T, Bertoletti L, Cucherat M, Vital-Durand D, et al. Double blind vs.
502 open design on treatment effect of new oral anticoagulants in atrial fibrillation: A meta-analysis.
503 *Fundamental and Clinical Pharmacology*. 2012;26(SUPPL.1):17.
- 504 59. Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study
505 design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern*
506 *Med*. 2012;157(6):429-38.
- 507 60. Unverzagt S, Prondzinsky R, Peinemann F. Single-center trials tend to provide larger treatment
508 effects than multicenter trials: a systematic review. *J Clin Epidemiol*. 2013;66(11):1271-80.
- 509 61. Hartling L, Hamm MP, Fernandes RM, Dryden DM, Vandermeer B. Quantifying bias in
510 randomized controlled trials in child health: a meta-epidemiological study. *PLoS One*. 2014;9(2):e88008.
- 511 62. Chaimani A, Vasiliadis HS, Pandis N, Schmid CH, Welton NJ, Salanti G. Effects of study precision
512 and risk of bias in networks of interventions: a network meta-epidemiological study. *Int J Epidemiol*.
513 2013;42(4):1120-31.
- 514 63. Bialy L, Vandermeer B, Lacaze-Masmonteil T, Dryden DM, Hartling L. A meta-epidemiological
515 study to examine the association between bias and treatment effects in neonatal trials. *Evid Based Child*
516 *Health*. 2014;9(4):1052-9.
- 517 64. Pildal J, Hróbjartsson A, Jørgensen KJ, Hilden J, Altman DG, Gøtzsche PC. Impact of allocation
518 concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol*.
519 2007;36(4):847-57.
- 520 65. Contopoulos-Ioannidis DG, Gilbody SM, Trikalinos TA, Churchill R, Wahlbeck K, Ioannidis JPA.
521 Comparison of Large Versus Smaller Randomized Trials for Mental Health-Related Interventions.
522 *American Journal of Psychiatry*. 2005;162(3):578-84.
- 523 66. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature
524 searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol*
525 *Assess*. 2003;7(1):1-76.
- 526 67. Siersma V, Als-Nielsen B, Chen W, Hilden J, Gluud LL, Gluud C. Multivariable modelling for meta-
527 epidemiological assessment of the association between trial quality and treatment effects estimated in
528 randomized clinical trials. *Stat Med*. 2007;26(14):2745-58.
- 529 68. Wood L, Egger M, Gluud LL, Schulz KF, Jüni P, Altman DG, et al. Empirical evidence of bias in
530 treatment effect estimates in controlled trials with different interventions and outcomes: meta-
531 epidemiological study. *Bmj*. 2008;336(7644):601-5.
- 532 69. Savović J, Turner RM, Mawdsley D, Jones HE, Beynon R, Higgins JPT, et al. Association Between
533 Risk-of-Bias Assessments and Results of Randomized Trials in Cochrane Reviews: The ROBES Meta-
534 Epidemiologic Study. *American Journal of Epidemiology*. 2017;187(5):1113-22.
- 535 70. Dechartres A, Altman DG, Trinquart L, Boutron I, Ravaud P. Association Between Analytic
536 Strategy and Estimates of Treatment Outcomes in Meta-analyses. *JAMA*. 2014;312(6):623-30.
- 537 71. van Tulder MW, Suttorp M, Morton S, Bouter LM, Shekelle P. Empirical evidence of an
538 association between internal validity and effect size in randomized controlled trials of low-back pain.
539 *Spine (Phila Pa 1976)*. 2009;34(16):1685-92.
- 540 72. Fenwick J, Needleman IG, Moles DR. The effect of bias on the magnitude of clinical outcomes in
541 periodontology: a pilot study. *J Clin Periodontol*. 2008;35(9):775-82.
- 542 73. Hempel S, Miles J, Suttorp MJ, Wang Z, Johnsen B, Morton S, et al. AHRQ Methods for Effective
543 Health Care. Detection of Associations Between Trial Quality and Effect Sizes. Rockville (MD): Agency for
544 Healthcare Research and Quality (US); 2012.

- 545 74. Welton NJ, Ades AE, Carlin JB, Altman DG, Sterne JAC. Models for potentially biased evidence in
546 meta-analysis using empirically based priors. *Journal of the Royal Statistical Society: Series A (Statistics*
547 *in Society)*. 2009;172(1):119-36.
- 548 75. Sterne JA, Jüni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the
549 influence of study characteristics on treatment effects in 'meta-epidemiological' research. *Stat Med*.
550 2002;21(11):1513-24.
- 551 76. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for
552 assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
- 553 77. Montori VM, Bhandari M, Devereaux PJ, Manns BJ, Ghali WA, Guyatt GH. In the dark: the
554 reporting of blinding status in randomized controlled trials. *J Clin Epidemiol*. 2002;55(8):787-90.
- 555 78. Akl EA, Sun X, Busse JW, Johnston BC, Briel M, Mulla S, et al. Specific instructions for estimating
556 unclearly reported blinding status in randomized trials were reliable and valid. *J Clin Epidemiol*.
557 2012;65(3):262-7.
- 558 79. Haahr MT, Hróbjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and
559 a survey of authors. *Clin Trials*. 2006;3(4):360-5.
- 560 80. Kent DM, Trikalinos TA. Therapeutic innovations, diminishing returns, and control rate
561 preservation. *Jama*. 2009;302(20):2254-6.
- 562 81. Pereira TV, Horwitz RJ, Ioannidis JP. Empirical evaluation of very large treatment effects of
563 medical interventions. *Jama*. 2012;308(16):1676-84.
- 564 82. Janiaud P, Agarwal A, Tzoulaki I, Theodoratou E, Tsilidis KK, Evangelou E, et al. Validity of
565 observational evidence on putative risk and protective factors: appraisal of 3744 meta-analyses on 57
566 topics. *BMC Med*. 2021;19(1):157.
- 567 83. Anglemeyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study
568 designs compared with those assessed in randomized trials. *Cochrane Database Syst Rev*.
569 2014;2014(4):Mr000034.
- 570 84. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N*
571 *Engl J Med*. 2000;342(25):1878-86.
- 572 85. Treweek S, Bevan S, Bower P, Campbell M, Christie J, Clarke M, et al. Trial Forge Guidance 1:
573 what is a Study Within A Trial (SWAT)? *Trials*. 2018;19(1):139.

574

Identification

Screening

Included

Records removed before screening
(n= 1397)

Records screened
(n= 6623)

Records excluded
(n= 6456)

Records retrieved in full
(n= 247)

Duplicates
(n= 80)

Full-text records assessed for eligibility
(n= 167)

Records excluded
Animal study (n= 3)
Duplicates (n= 2)
No blinding assessment (n= 23)
No assessment of treatment effects (n= 14)
No comparison (n= 37)
No full text or abstract (n= 29)
Insufficient data (n= 4)
Intervention/Outcome unclear (n= 3)
Improper study type (n= 5)

Records included in review
(n= 47)

