Meta-analyses frequently include old trials that are associated with a larger intervention effect: a meta-epidemiological study

Violaine Smail-Faugeron, Aidan Tan, Agnès Caille, Youri Yordanov, David Hajage, Florence Tubach, Guillaume Martin, Agnès Dechartres

Meta-analyses frequently include old trials that are associated with a larger intervention effect: a meta-epidemiological study

Violaine SMAIL-FAUGERON, DDS, PhD[1], Aidan TAN, MD, MPH[2], Agnès CAILLE, MD, PhD[1,3*], Youri YORDANOV, MD, PhD[4*], David HAJAGE, MD, PhD[1], Florence TUBACH, MD, PhD [1], Guillaume MARTIN, MD[1], Agnès DECHARTRES, MD, PhD [1]

[1] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpital Pitié-Salpêtrière, Département de Santé Publique, Paris, France

[2] Health Services and Outcomes Research, National Healthcare Group, Singapore, Singapore

[3] Université de Tours, Université de Nantes, INSERM, SPHERE U1246, Tours, France; INSERM CIC 1415, CHRU de Tours, Tours, France

[4] Sorbonne Université, INSERM, Institut Pierre Louis d'Epidémiologie et de Santé Publique, AP-HP, Hôpital Saint-Antoine, Service d'Accueil des Urgences, Paris, France

* AC and YY contributed equally to this article

Correspondance to:

Agnes Dechartres, MD, PhD

Sorbonne Université, INSERM U1136, équipe PEPITES, Institut Pierre Louis d'épidémiologie et de santé publique

Hôpital Pitié-Salpêtrière, 47-83 Boulevard de l'Hôpital 75013 Paris, France

Email: agnes.dechartres@aphp.fr

Phone: (33) 1 42 16 05 99

Word count: 3386

Key-words: systematic review, meta-analysis, meta-epidemiology, external validity, publication date

ABSTRACT (199 words)

Objective: To assess whether meta-analyses include older randomized controlled trials (RCTs) and whether intervention effect differ between older and recent RCTs.

Study design and setting: In this meta-epidemiological study of 295 meta-analyses (2940 RCTs) published in 2017-2018, we evaluated the difference in intervention effects between older (ie, published before 2000) and recent RCTs. We also compared effects by quarters of publication year within each meta-analysis (from quarter 1 including the 25% oldest trials to quarter 4 including the 25% most recent trials). A ratio of odds ratio (ROR) < 1 indicates larger effects in older than recent RCTs.

Results: Trials published before 2000 and before 1990 represented 25% and 10% of all trials, respectively. Intervention effects were significantly larger for old than recent RCTs (ROR=0.92, 95% confidence interval[CI] 0.85 to 1.00, I^2=22%). Compared with the most recent trials (quarter 4), intervention effects were significantly larger for the oldest trials (quarter 1) (ROR=0.85, 95% CI 0.79 to 0.92) and for trials in quarter 2 (ROR=0.89, 95% CI 0.83 to 0.96) but not for trials in quarter 3 (ROR=0.98, 95% CI 0.91 to 1.05).

Conclusions: Intervention effects were larger for older than recent RCTs. Meta-analyses including older trials should be interpreted cautiously.

Key-words: systematic review, meta-analysis, meta-epidemiology, external validity, publication date

Running title: Inclusion of old trials in meta-analyses

What is new?

Key findings

- Trials published before 2000 represented one fourth of all trials included in meta-analyses and trials published before 1990 almost 10%. The oldest trial was published in 1951.
- Intervention effects were, on average, significantly larger for older than recent trials.
- Results were consistent in sensitivity analyses adjusted on risk of bias and sample size.

What this adds to what was known?

- It is generally recommended to include all trials within a meta-analysis whatever their publication date but this may have an influence on external validity and intervention effect.

What is the implication and what should change now?

- With a biomedical literature that is increasing exponentially, we wonder whether it is reasonable to consider the results of old trials sometimes conducted more than 50 years ago.

INTRODUCTION

Systematic reviews and meta-analyses of randomized controlled trials (RCTs) are usually considered to provide the highest level of evidence for the efficacy of medical interventions and have become essential tools for healthcare decision-making(1-4).

A key question when conducting a systematic review is whether to search for all available evidence regardless of publication date or to restrict the search to recent studies. Cochrane generally recommends against date restrictions, except "if it is known that relevant studies could only have been reported during a specific time period, for example if the intervention was only available after a certain time point"(5).

Therefore, meta-analyses may include very old trials, sometimes performed up to 50 years ago. Others are based on old trials only and do not include any recent trials. A study published in 2009 reported that among a sample of 157 meta-analyses, 30% included no trial published in the preceding 10 years and only 8% discussed the potential consequences of including older trials(6).

Including all RCTs in a meta-analysis regardless of their publication date raises potential concerns. Major changes may have occurred over time in clinical practice and patient management, in the organization of health care systems but also in trial characteristics, with an improvement of their methods over time(7). All these changes may limit the generalizability of older trial results(8) and may affect the intervention effect itself(9, 10). This is also inconsistent with the evidence-based medicine principle that recommends focusing on the current best evidence(3).

In this meta-epidemiological study, we aimed to assess whether recent systematic reviews with meta-analyses included older trials and whether intervention effect estimates differed between older and recent RCTs.

METHODS

Search strategy

We searched PubMed for all systematic reviews published from January 2017 to October 2018 in the Cochrane Database of Systematic Reviews and in the 5 journals with the highest impact factors within each medical category according to the *Journal Citation Reports* (Appendix 1).

Study selection

Selection of systematic reviews

We included systematic reviews of health-related interventions with at least one meta-analysis of three or more RCTs. We excluded network or Bayesian meta-analyses to have a more homogeneous sample, cost-effectiveness analyses, methodological reviews, overviews, protocols, withdrawn reviews, viewpoints, as well as meta-analyses of individual patient data or including non-randomized or animal studies.

Selection of meta-analyses

From each eligible review, we selected all comparisons of an experimental intervention to no intervention, placebo or usual care because of the uncertainty in the direction of bias for comparisons of two active interventions.

From each eligible comparison, we selected a single meta-analysis of a binary outcome. In cases of multiple eligible outcomes, we chose the primary outcome or the first one presented. We did not consider meta-analyses of adverse events because the direction of bias is uncertain in this case. If results were presented as a subgroup analysis, we selected the meta-analysis only if there was an overall estimate of intervention effect.

We identified overlapping meta-analyses sharing at least one RCT and excluded the meta-analysis that included fewer trials(11).

Two authors (VSF and AT) independently conducted the selection process. Any disagreements were resolved by discussion with the help of a third reviewer (AD) whenever necessary to achieve consensus.

Data extraction

Using a standardized data extraction form, two authors (VSF and AT) independently extracted the following characteristics from the review report, with discrepancies resolved by discussion with a third reviewer (AD) whenever necessary to achieve consensus.

Systematic review and meta-analysis characteristics

- General characteristics: publication year reported in the systematic review, journal, medical specialty, funding sources
- Year of literature search and whether the authors reported a restriction by date
- Interventions assessed
- Outcome evaluated and whether it was objective or subjective(12)
- Number of RCTs included in the meta-analysis
- Range of publication years of RCTs included in the meta-analysis
- Reporting of subgroup or sensitivity analyses based on RCT publication year
- Whether and how authors discussed the publication range of RCTs in the Discussion section of the review.

Trial characteristics

- General characteristics: first author, publication year reported in the systematic review, funding sources
- Recruitment start and end dates
- Risk of bias assessment for the following items of the Cochrane Risk of Bias tool: sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment and incomplete outcome data.

- Results: number of patients analyzed and number of events for the experimental and control groups, or risk ratios (RRs) or odds ratios (ORs) and the 95% confidence intervals (CIs).

Definition of older trials

Because the definition of "older" trial may vary across research questions, we used different approaches.

Pre-specified cutoffs

First, we considered older trials as those published before the year 2000. We chose this year because the Consolidated Standards of Reporting Trials (CONSORT) Statement first published in 1996 improved the reporting of clinical trials and probably their methodological quality(13-18). We chose 2000 rather than 1996 to account for the trial recruitment period and implementation of the CONSORT Statement. We also considered quarters of publication year within each meta-analysis from quarter 1 including the 25% oldest trials to quarter 4 including the 25% most recent trials.

Evaluation of change in clinical practice

We also defined older trials as those conducted before an important change in clinical practice. To identify a possible change, for each meta-analysis, we contacted the corresponding author of the review to ask whether they had identified a substantial change in clinical practice while conducting the review. We also systematically contacted the corresponding author of the two most recent RCTs to ask about a potential change in clinical practice over time and when it occurred. We also asked RCT authors if they had to conduct a meta-analysis on the topic, whether they would restrict the search by date and if yes, when.

Data synthesis

For each RCT, we derived an intervention effect estimate expressed as ORs and associated variance. Outcomes were re-coded whenever necessary so that an OR <1 indicated a

beneficial effect of the experimental intervention. In cases of 0-count cells in one group, we used the continuity correction proposed by Sweeting et al.(19). We estimated the difference in intervention effect estimates between older and recent RCTs by a two-step meta-epidemiological method(20). For each meta-analysis including at least one older and one recent trial, we estimated the ratio of intervention-effect OR (ROR) for older RCTs to that for recent RCTs with a random-effects meta-regression model. We then estimated a combined ROR and 95% CIs across meta-analyses by a random-effects meta-analysis model to account for heterogeneity among meta-analyses. An ROR < 1 indicates that older RCTs yield larger estimates than recent RCTs. Heterogeneity in RORs across meta-analyses was assessed by the heterogeneity test, I^2 statistic and the between–meta-analyses variance τ^2.

As a secondary analysis, we compared intervention effects by quarters of publication year within the same meta-analysis by using a multilevel logistic regression model with random effects(21) (Appendix 2).

Subgroup and sensitivity analyses

We performed a subgroup analysis to compare the difference in intervention effects between older and recent RCTs in Cochrane and non-Cochrane meta-analyses.

Sensitivity analyses involved adjusting the meta-regression model for each domain of the Risk of Bias tool(22-26) and sample size(27).

Comparison of intervention effects between the overall meta-analysis and meta-analysis restricted to trials published in the last 10 years

For meta-analyses including at least one trial published in the 10 years preceding the review publication, we repeated the meta-analysis but including only these trials. We then compared the results and statistical significance with the meta-analysis including all trials.

Meta-analyses conducted before and after a practice change

For meta-analyses for which review or RCT authors indicated a change in clinical practice, we performed a meta-analysis of RCTs conducted before this change and a meta-analysis of RCTs conducted after this change based on RCT recruitment period when the recruitment period was reported.

Analyses involved using R 4.0.2 (http://www.R-project.org, the R Foundation for Statistical Computing, Vienna, Austria) and SAS 9.4 (SAS Institute Inc.).

RESULTS

Of 3289 references identified by the search, 287 systematic reviews were selected (Appendix Figure 1). There was a total of 295 meta-analyses including 2940 RCTs.

Characteristics of included meta-analyses

Of the 295 meta-analyses, 224 (75.9%) were from Cochrane reviews and 71 (24.1%) from non-Cochrane reviews. The most represented medical condition was obstetrics for Cochrane meta-analyses (n=41, 18.3%) and anesthesia/critical care for non-Cochrane meta-analyses (n=11, 15.5%). The search was recent for most reviews: 219 (97.8%) Cochrane reviews and 57 (80.3%) non-Cochrane reviews had a search date within the last 2 years. A pharmacological intervention was evaluated in 135 (60.3%) and 43 (60.6%) Cochrane and non-Cochrane meta-analyses. The median number of RCTs per meta-analysis was 6 (IQR, 4 to 11) and 9 (IQR, 5 to 15), respectively (Appendix Table 1).

Inclusion of older trials in meta-analyses

None of the reviews reported a restriction by date when conducting the search. Cochrane meta-analyses included more frequently than non-Cochrane meta-analyses at least one RCT published before 2000 (130 [58.0%] vs. 26 [36.6%]) and one RCT published before 1980 (29 [12.9%] vs. 2 [2.8%]). Overall, trials published before 2000 represented one quarter of all trials (n=745) and trials published before 1990, almost 10% (n=268). The oldest trial was published in 1951. The median age of all trials was 12 (IQR 6 to 21) and 6 (IQR 3 to 11) years in Cochrane and non-Cochrane meta-analyses (Figure 1) and the median age of the oldest trial was 20 (IQR 13 to 31) and 14 (IQR 10 to 20) years. Some meta-analyses, all Cochrane, were based on older trials only: 44 (19.6%) included no trial published in the preceding 10 years of publication, 21 (9.4%) had all trials published before 2000 and two (0.9%) had all trials published before 1990. Only two Cochrane reviews performed subgroup or sensitivity analyses based on RCT publication year. Four other Cochrane reviews planned

to do so but had insufficient data, and another one discussed year of publication of the included RCTs. Of the 224 Cochrane reviews (all using GRADE to rate the certainty of evidence), only four downgraded the quality of evidence because of inclusion of older trials.

Characteristics of RCTs published before and after 2000

The sample size was smaller in older than recent trials (median sample size=93 [IQR 48 to 227] vs. 133 [IQR 68 to 360]). Older trials were more frequently at high or unclear risk of bias than recent trials for sequence generation, blinding of participants and personnel, blinding of outcome assessors, and incomplete outcome data. The median number of items at high or unclear risk was 3 (IQR 2 to 4) for older trials and 2 (IQR 0 to 3) for recent trials (Table 1).

Evaluation of change in clinical practice

We contacted 861 authors: 287 authors of systematic reviews and 590 authors of RCTs. In total, 125 (43.6%) review authors and 116 (19.7%) trialists answered. Most review authors (n=71, 56.8%) did not identify a substantial change in clinical practice, 11 (8.8%) identified one change and 43 (34.4%) had no opinion or did not answer the question. Among the 116 trialists, 57 (49.1%) did not identify a substantial change in clinical practice, 41 (35.3%) identified one change and 18 (15.6%) had no opinion or did not answer the question. There were 38 meta-analyses for which both review authors and trialists answered: for 20 meta-analyses, review authors did not identify a substantial change in clinical practice, whereas trialists identified one change; for two meta-analyses, it was the opposite; and for the last 16, both author types did not identify a substantial change in clinical practice.

Of the 98 trialists answering the question whether they would apply a restriction if they had to conduct a meta-analysis on the topic, 49 (50%) answered no, including three who identified a change in clinical practice. Seven (7.1%) would include all studies but would perform a

subgroup analysis by time and 42 (42.9%) would apply a restriction. Among these 42, most considered 2000 (n=22, 52.4%) or the last 10 to 15 years (n=14, 33.3%) as a relevant cutoff.

Difference in intervention effect estimates between older and recent RCTs

A total of 135 meta-analyses (1632 RCTs) included both RCTs published before and after 2000. The intervention effect estimates were, on average, significantly larger in older than recent RCTs (combined ROR 0.92, 95% CI, 0.85 to 1.00) (Figure 2). There was some heterogeneity across individual meta-analyses (p=0.02, I^2 = 22%, τ^2 = 0.0350).

In the secondary analysis comparing quarters of publication year within meta-analyses, intervention effect estimates were 15% larger for the oldest trials (quarter 1); ROR 0.85, 95% CI 0.79 to 0.92), 10% larger for trials in quarter 2 (ROR 0.89, 95% CI 0.83 to 0.96), and not significantly larger for trials in quarter 3 (ROR 0.98, 95% CI 0.91 to 1.05) as compared with the most recent trials (quarter 4) (Figure 3).

Subgroup and sensitivity analyses

We found no evidence of a difference between Cochrane and non-Cochrane meta-analyses (ROR 0.92, 95% CI 0.85 to 1.00, I^2=15% for Cochrane meta-analyses and ROR 0.93, 95% CI 0.75 to 1.15, I^2=45% for non-Cochrane meta-analyses, $p_{interaction}$ = 0.91). We found consistent results in sensitivity analyses (Figure 4).

Effect of restricting meta-analyses to trials published in the last 10 years

When comparing results between the overall meta-analysis and the meta-analysis restricted to trials published in the last 10 years, statistical significance was changed from non-significant to significant or from significant to non-significant in 24 (16.1%) of the 149 meta-analyses including at least one trial published the last 10 years. Results were no longer significant when restricting to trials published in the last 10 years for 17 (11.4%) meta-analyses but were significant and not the overall meta-analysis for 7 (4.7%) (Appendix Figure 2).

Meta-analyses conducted before and after a change in clinical practice

The meta-analyses of RCTs conducted before and after a change in clinical practice identified by RCT and review authors is shown in Appendix Figure 3, but this analysis was based on 12 meta-analyses for which the authors identified a change in practice and whose study recruitment period was reported and was before or after the practice change. When comparing results between meta-analyses of RCTs published before and after a practice change, statistical significance was changed from non-significant to significant in 3 (25%) and from significant to non-significant in one (8%).

DISCUSSION

In this meta-epidemiological study, we comprehensively studied to what extent older trials were included in recent meta-analyses published by Cochrane and high impact factor journals and their association with intervention effects. We also attempted to identify a possible change in clinical practice by contacting RCT and review authors. None of the included reviews applied a restriction by date and many included older trials. Older RCTs showed significantly larger intervention effects on average than did recent RCTs.

We found a high proportion of older trials in some meta-analyses, especially Cochrane reviews with more than half including at least one RCT published before 2000. Cochrane reviews are regularly updated, adding new trials to older ones that remain in updated reviews, which may explain this finding.

Intervention effects were slightly larger in older than recent RCTs, which is consistent with studies suggesting that intervention effect estimates reported in RCTs are decreasing over time(28-33). Several reasons may explain this finding. First, improvement in standard of care over time might lead to reduced effects of experimental interventions particularly if components of the active intervention become part of standard care. Another reason is related to the higher risk of bias in older than recent trials. A high or unclear risk of bias, particularly for sequence generation and allocation concealment, was previously found associated with an overestimation of intervention effects(22-26, 34). The larger intervention effect in older trials may also be related to the small-study effect, defined as the tendency for smaller studies to show larger intervention effect than larger studies in a meta-analysis(35) in that we found that older RCTs had smaller sample sizes than did recent RCTs. However, we found consistent results in sensitivity analyses.

We would have expected a larger difference in intervention effect estimates between older and recent RCTs. We found some heterogeneity across meta-analyses and also significantly

larger intervention effects in recent than older RCTs for four meta-analyses. Seven meta-analyses had significant results when restricting to recent trials, but no significant effect when including all trials. The characteristics of these 11 meta-analyses are in Appendix Table 2. This finding suggests that the difference in intervention effects between older and recent trials may go in the opposite direction for some meta-analyses, which may explain in part why the combined ROR was not larger. Moreover, we considered the same pre-specified cutoff — 2000 — for all meta-analyses, regardless of the topic, but this cut-off may not be appropriate for all topics, which may have resulted in an attenuation of the average difference. This is why we also considered quarters of publication year within each meta-analysis. We also evaluated possible changes in practice for each topic. Because we could not identify experts of each domain, we contacted authors of the two most recent RCTs as well as review authors. Despite a low response rate, more trialists than review authors considered that a substantial change occurred over time and would apply a restriction by date if they had to perform a meta-analysis. Among those, half of trialists considered 2000 as a relevant cutoff for meta-analyses, which is consistent with our pre-specified cutoff. Nevertheless, other trialists considered an even more recent cutoff.

Our study has important implications for the planning and conduct of systematic reviews. In particular, our results raise the question of whether all available evidence should be considered. We found that including all trials whatever their publication date may in general lead to an overestimation of intervention effects estimates and sometimes an underestimation, with a possible impact on meta-analysis conclusions. It seems more appropriate to focus on more recent RCTs because they are more likely to have a larger sample size and lower risk of bias and be more representative of current practice. This is also consistent with the evidence-based medicine principle that recommends focusing on the current best evidence(3). Moreover, including older trials, that are no longer relevant to current clinical practices, may

affect the similarity of the studies included in a meta-analysis and not only effect size). However, fixing a relevant cutoff is difficult. Two approaches may be discussed: 1) pre-specifying the same cutoff, for example, 2000 for all meta-analyses regardless of the topicor 2) defining a specific cutoff by topic. The first approach is probably not appropriate for all topics because an important change in clinical practice may occur later for some of them. The second approach is personalized however, the risk of introducing bias is real if the cutoff is based on knowledge of the literature so review authors would need to well-justify the cutoff. At a minimum, meta-analysts should be encouraged to perform sensitivity analyses to evaluate the evolution of intervention effects over time and to assess whether the results of the overall meta-analysis agree with the results of the most recent trials.

Our study has limitations. First, the search strategy might have missed some meta-analyses. We selected meta-analyses published in the Cochrane Database of Systematic Reviews and in the leading journals of each medical specialty because they are more likely to publish reviews of broad medical interest and to be better reported(7) but studies should not be judged based on the journal in which they are published. We extracted characteristics of RCTs from review reports, and most did not report the recruitment period or funding sources. This is why we relied on publication date rather than recruitment period. We cannot exclude a risk of confounding. We adjusted for risk of bias and sample size. Other characteristics may be additional confounders such as the trial funding source. Finally, we had a very low response rate from authors limiting the evaluation of change in clinical practice over time.

In conclusion, we found that older trials may yield a different intervention effect than recent trials: more frequently a larger effect and sometimes a lower effect. With a biomedical literature that is increasing exponentially, we wonder whether it is reasonable to consider in a meta-analysis the results of old trials sometimes conducted more than 50 years ago as this may not reflect current practice and have an impact on intervention effect.

ACKNOWLEDGEMENTS

Laura Smales, BioMedEditing, for English editing of the manuscript.

FUNDING SUPPORT

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

COMPETING INTERESTS

The authors declare no competing interests.

AUTHORS' CONTRIBUTIONS

Violaine Smaïl-Faugeron participated in the design of the study, selection and acquisition of data, participated to statistical analyses, interpreted the data and wrote the manuscript.

Aidan Tan was involved in the selection and acquisition of data and critically reviewed the manuscript.

Agnes Caille participated to statistical analysis and critically reviewed the manuscript.

Youri Yordanov was in charge of the survey and critically reviewed the manuscript.

David Hajage critically reviewed the manuscript and help with designing figures.

Florence Tubach critically reviewed the manuscript.

Guillaume Martin conducted statistical analysis and critically reviewed the manuscript.

Agnès Dechartres participated in the design of the study, interpretation of data and writing of the manuscript.

DATA SHARING STATEMENT

Raw data and analysis available on request from the corresponding author.

REFERENCES

1. Chalmers I, Altman D. Systematic Reviews. London: BMJ. 1995.
2. Djulbegovic B, Guyatt GH. Progress in evidence-based medicine: a quarter century on. Lancet. 2017;390(10092):415-23.
3. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. Bmj. 1996;312(7023):71-2.
4. Sox HC. Defining comparative effectiveness research: the importance of getting it right. Med Care. 2010;48(6 Suppl):S7-8.
5. Higgins J, Green S, (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.handbook.cochrane.org.
6. Patsopoulos NA, Ioannidis JP. The use of older studies in meta-analyses of medical interventions: a survey. Open Med. 2009;3(2):e62-8.
7. Dechartres A, Trinquart L, Atal I, Moher D, Dickersin K, Boutron I, et al. Evolution of poor reporting and inadequate methods over time in 20 920 randomised controlled trials included in Cochrane reviews: research on research study. Bmj. 2017;357:j2490.
8. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". Lancet. 2005;365(9453):82-93.
9. Barroso J, Sandelowski M, Voils CI. Research results have expiration dates: ensuring timely systematic reviews. J Eval Clin Pract. 2006;12(4):454-62.
10. Gurevitch J, Koricheva J, Nakagawa S, Stewart G. Meta-analysis and the science of research synthesis. Nature. 2018;555(7695):175-82.
11. Savovic J, Harris RJ, Wood L, Beynon R, Altman D, Als-Nielsen B, et al. Development of a combined database for meta-epidemiological research. Res Synth Methods. 2011;2(1):78.
12. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429-38.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. Jama. 1996;276(8):637-9.
14. Cobo E, Cortes J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: masked randomised trial. Bmj. 2011;343:d6783.
15. Hopewell S, Altman DG, Moher D, Schulz KF. Endorsement of the CONSORT Statement by high impact factor medical journals: a survey of journal editors and journal 'Instructions to Authors'. Trials. 2008;9:20.
16. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. Bmj. 2012;344:e4178.
17. Shamseer L, Hopewell S, Altman DG, Moher D, Schulz KF. Update on the endorsement of CONSORT by high impact factor journals: a survey of journal "Instructions to Authors" in 2014. Trials. 2016;17(1):301.
18. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. Cochrane Database Syst Rev. 2012;11:Mr000030.

19. Sweeting MJ, Sutton AJ, Lambert PC. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of sparse data. Stat Med. 2004;23(9):1351-75.
20. Sterne JA, Juni P, Schulz KF, Altman DG, Bartlett C, Egger M. Statistical methods for assessing the influence of study characteristics on treatment effects in 'meta-epidemiological' research. Stat Med. 2002;21(11):1513-24.
21. Siersma V, Als-Nielsen B, Chen W, Hilden J, Gluud LL, Gluud C. Multivariable modelling for meta-epidemiological assessment of the association between trial quality and treatment effects estimated in randomized clinical trials. Stat Med. 2007;26(14):2745-58.
22. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet. 1998;352(9128):609-13.
23. Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Burgi E, Scherer M, et al. The effects of excluding patients from the analysis in randomised controlled trials: meta-epidemiological study. Bmj. 2009;339:b3244.
24. Pildal J, Hrobjartsson A, Jorgensen KJ, Hilden J, Altman DG, Gotzsche PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. Int J Epidemiol. 2007;36(4):847-57.
25. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. Jama. 1995;273(5):408-12.
26. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. Bmj. 2008;336(7644):601-5.
27. Dechartres A, Trinquart L, Boutron I, Ravaud P. Influence of trial sample size on treatment effect estimates: meta-epidemiological study. Bmj. 2013;346:f2304.
28. Gehr BT, Weiss C, Porzsolt F. The fading of reported effectiveness. A meta-analysis of randomised controlled trials. BMC Med Res Methodol. 2006;6:25.
29. Ioannidis J, Lau J. Evolution of treatment effects over time: empirical insight from recursive cumulative metaanalyses. Proc Natl Acad Sci U S A. 2001;98(3):831-6.
30. Ioannidis JP. Contradicted and initially stronger effects in highly cited clinical research. Jama. 2005;294(2):218-28.
31. Trikalinos TA, Churchill R, Ferri M, Leucht S, Tuunainen A, Wahlbeck K, et al. Effect sizes in cumulative meta-analyses of mental health randomized trials evolved over time. J Clin Epidemiol. 2004;57(11):1124-30.
32. Alahdab F, Farah W, Almasri J, Barrionuevo P, Zaiem F, Benkhadra R, et al. Treatment Effect in Earlier Trials of Patients With Chronic Medical Conditions: A Meta-Epidemiologic Study. Mayo Clin Proc. 2018;93(3):278-83.
33. Wang Z, Alahdab F, Almasri J, Haydour Q, Mohammed K, Abu Dabrh AM, et al. Early studies reported extreme findings with large variability: a meta-epidemiologic study in the field of endocrinology. J Clin Epidemiol. 2016;72:27-32.
34. Tierney JF, Stewart LA. Investigating patient exclusion bias in meta-analysis. Int J Epidemiol. 2005;34(1):79-87.
35. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol. 2000;53(11):1119-29.

FIGURE LEGENDS

Figure 1. Age of randomized controlled trials in Cochrane and non-Cochrane meta-analyses.

Legend: The box is drawn from Q1 to Q3, the horizontal line in the box indicates the median value. Q1: Quartile 1, Q3: Quartile 3

Figure 2. Difference in intervention effect estimates between older and recent randomized controlled trials (RCTs) by meta-epidemiological analysis.

Legend: This analysis is based on 135 meta-analyses including at least one older and one recent trial for a total of 1632 RCTs.

Figure 3. Comparison of intervention effect estimates by trial publication year, grouped by quarters (from quarter [Q] 1 with the oldest trials, to Q4 with the most recent trials)

Legend: This analysis is based on 244 meta-analyses (2748 RCTs).

Figure 4. Sensitivity analyses with adjustment on each domain of the Cochrane Risk of Bias tool and sample size

Legend: ROR= ratio of odds ratio

Appendix Figure 1. Flow chart of the selection process for Cochrane reviews and non-Cochrane reviews

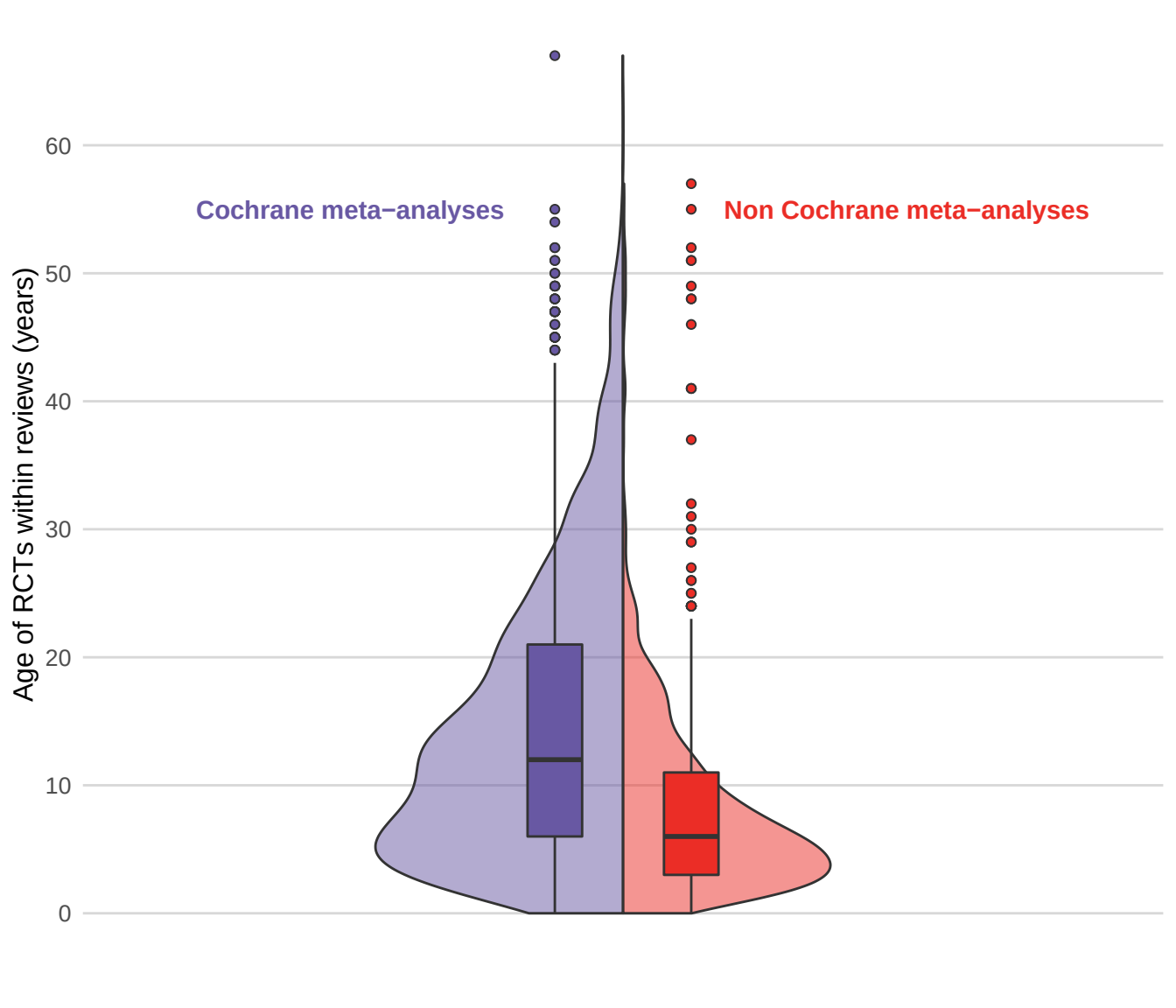
Appendix Figure 2. Effect on statistical significance of restricting the meta-analysis to trials published in the preceding 10 years

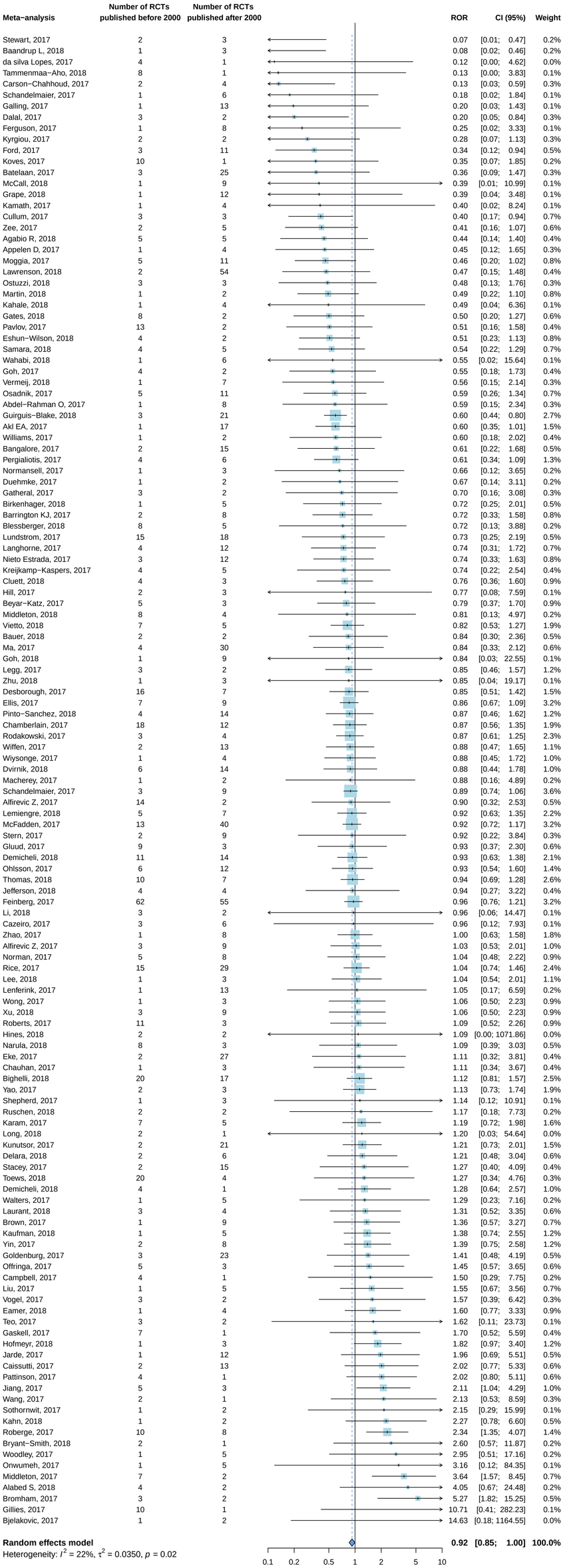
Appendix Figure 3. Meta-analyses of randomized controlled trials conducted before and after a practice change identified by randomized controlled trial authors or review authors

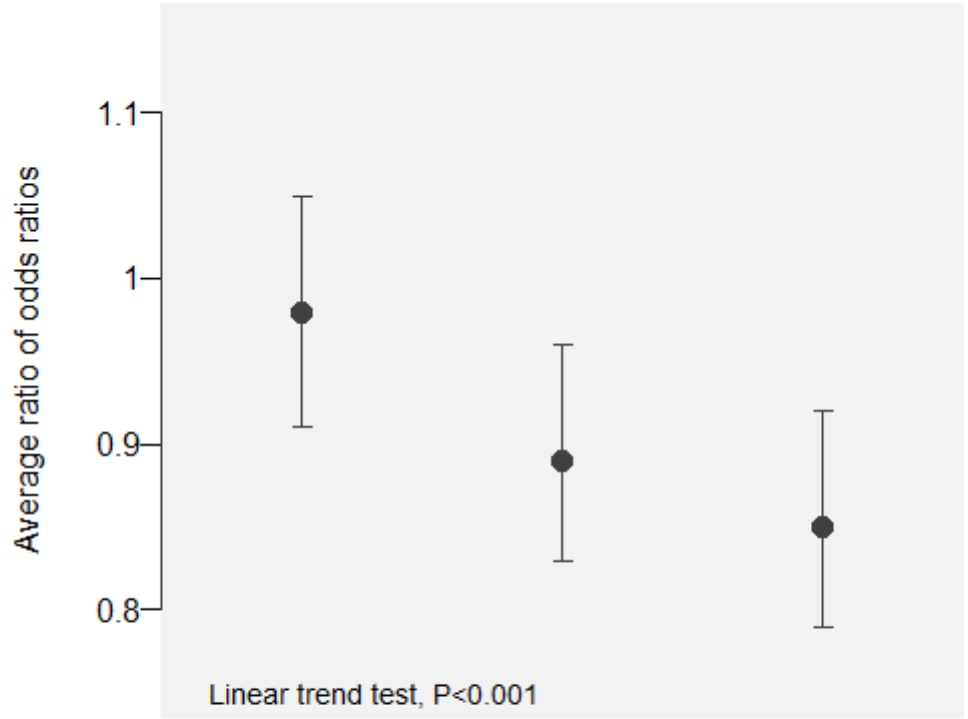
Table 1. Characteristics of older and recent randomized controlled trials. Values are number (percentage) unless otherwise indicated.

Characteristic	Older trials[§] (N=745)	Recent trials[§] (N=2195)
Funding source		
Public	81 (10.9)	307 (14.0)
Private	103 (13.8)	418 (19.0)
Both public and private	24 (3.2)	52 (2.4)
Not specific funding	6 (0.8)	62 (2.8)
Not reported in the review	531 (71.3)	1356 (61.8)
Intervention		
Pharmacologic	452 (60.7)	1349 (61.5)
Non-pharmacologic	293 (39.3)	846 (38.5)
Sample size (median (IQR, range))	93 (48–227, 5–110,150)	133 (68–360, 7–3,948,572)
Sample size		
<100 patients	382 (51.3)	868 (39.5)
100-200 patients	157 (21.1)	481 (21.9)
>200 patients	206 (27.6)	846 (38.6)
Cochrane Risk of Bias tool		
Sequence generation		
High	38 (5.1)	67 (3.1)
Low	235 (31.5)	1137 (51.8)
Unclear	427 (57.3)	683 (31.1)
Not reported in the review	45 (6.1)	308 (14.0)
Allocation concealment		
High	40 (5.4)	82 (3.7)
Low	196 (26.3)	936 (42.6)
Unclear	456 (61.2)	879 (40.1)
Not reported in the review	53 (7.1)	298 (13.6)
Blinding of participants and personnel		
High	205 (27.6)	690 (31.4)
Low	231 (31.0)	682 (31.1)
Unclear	235 (31.5)	459 (20.9)
Not reported in the review	74 (9.9)	364 (16.6)
Blinding of outcome assessors		
High	86 (11.5)	260 (11.8)
Low	248 (33.3)	889 (40.5)
Unclear	350 (47.0)	684 (31.2)
Not reported in the review	61 (8.2)	362 (16.5)
Incomplete outcome data		
High	131 (17.6)	253 (11.6)
Low	346 (46.4)	1276 (58.1)
Unclear	213 (28.6)	356 (16.2)
Not reported in the review	55 (7.4)	310 (14.1)
Number of items considered at high/unclear risk (median (IQR, range))	3 (2–4, 0–5)	2 (0–3, 0–5)

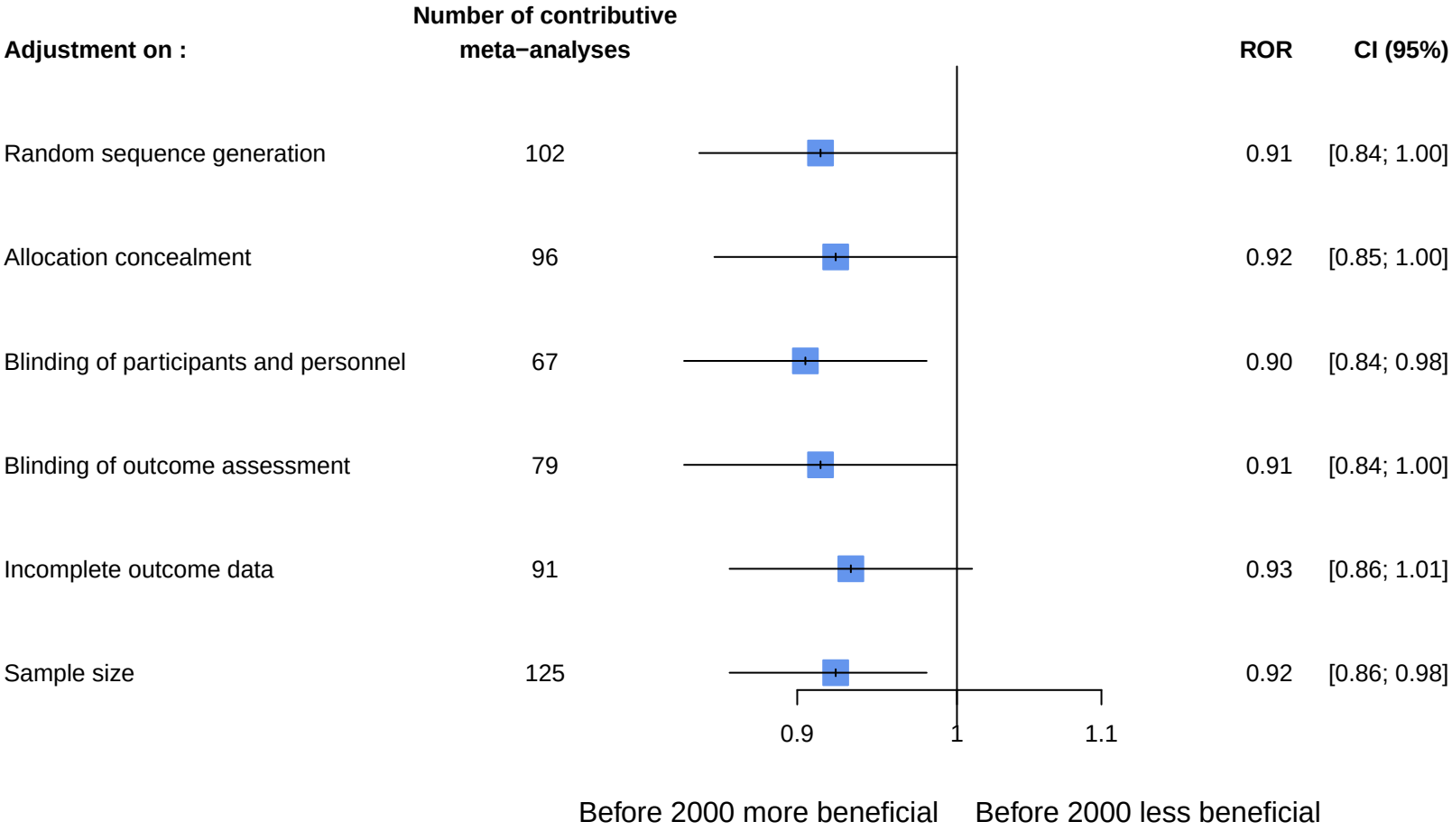
[§] Older trials are trials published before 2000, recent trials are trials published in 2000 or after.
IQR=interquartile range; RCT=randomized controlled trial


Cochrane meta−analyses
Non Cochrane meta−analyses
Age of RCTs within reviews (years)
60
50
40
30
20
10
0


Meta−analysis
Number of RCTs published before 2000
Number of RCTs published after 2000
ROR
CI (95%)
Weight
Stewart, 2017 2 3 0.07 [0.01; 0.47] 0.2%
Baandrup L, 2018 1 3 0.08 [0.02; 0.46] 0.2%
da silva Lopes, 2017 4 1 0.12 [0.00; 4.62] 0.0%
Tammenmaa−Aho, 2018 8 1 0.13 [0.00; 3.83] 0.1%
Carson−Chahhoud, 2017 2 4 0.13 [0.03; 0.59] 0.3%
Schandelmaier, 2017 1 6 0.18 [0.02; 1.84] 0.1%
Galling, 2017 1 13 0.20 [0.03; 1.43] 0.1%
Dalal, 2017 3 2 0.20 [0.05; 0.84] 0.3%
Ferguson, 2017 1 8 0.25 [0.02; 3.33] 0.1%
Kyrgiou, 2017 2 2 0.28 [0.07; 1.13] 0.3%
Ford, 2017 3 11 0.34 [0.12; 0.94] 0.5%
Koves, 2017 10 1 0.35 [0.07; 1.85] 0.2%
Batelaan, 2017 3 25 0.36 [0.09; 1.47] 0.3%
McCall, 2018 1 9 0.39 [0.01; 10.99] 0.1%
Grape, 2018 1 12 0.39 [0.04; 3.48] 0.1%
Kamath, 2017 1 4 0.40 [0.02; 8.24] 0.1%
Cullum, 2017 3 3 0.40 [0.17; 0.94] 0.7%
Zee, 2017 2 5 0.41 [0.16; 1.07] 0.6%
Agabio R, 2018 5 5 0.44 [0.14; 1.40] 0.4%
Appelen D, 2017 1 4 0.45 [0.12; 1.65] 0.3%
Moggia, 2017 5 11 0.46 [0.20; 1.02] 0.8%
Lawrenson, 2018 2 54 0.47 [0.15; 1.48] 0.4%
Ostuzzi, 2018 3 3 0.48 [0.13; 1.76] 0.3%
Martin, 2018 1 2 0.49 [0.22; 1.10] 0.8%
Kahale, 2018 1 4 0.49 [0.04; 6.36] 0.1%
Gates, 2018 8 2 0.50 [0.20; 1.27] 0.6%
Pavlov, 2017 13 2 0.51 [0.16; 1.58] 0.4%
Eshun−Wilson, 2018 4 2 0.51 [0.23; 1.13] 0.8%
Samara, 2018 4 5 0.54 [0.22; 1.29] 0.7%
Wahabi, 2018 1 6 0.55 [0.02; 15.64] 0.1%
Goh, 2017 4 2 0.55 [0.18; 1.73] 0.4%
Vermeij, 2018 1 7 0.56 [0.15; 2.14] 0.3%
Osadnik, 2017 5 11 0.59 [0.26; 1.34] 0.7%
Abdel−Rahman O, 2017 1 8 0.59 [0.15; 2.34] 0.3%
Guirguis−Blake, 2018 3 21 0.60 [0.44; 0.80] 2.7%
Akl EA, 2017 1 17 0.60 [0.35; 1.01] 1.5%
Williams, 2017 1 2 0.60 [0.18; 2.02] 0.4%
Bangalore, 2017 2 15 0.61 [0.22; 1.68] 0.5%
Pergialiotis, 2017 4 6 0.61 [0.34; 1.09] 1.3%
Normansell, 2017 1 3 0.66 [0.12; 3.65] 0.2%
Duehmke, 2017 1 2 0.67 [0.14; 3.11] 0.2%
Gatheral, 2017 3 2 0.70 [0.16; 3.08] 0.3%
Birkenhager, 2018 1 5 0.72 [0.25; 2.01] 0.5%
Barrington KJ, 2017 2 8 0.72 [0.33; 1.58] 0.8%
Blessberger, 2018 8 5 0.72 [0.13; 3.88] 0.2%
Lundstrom, 2017 15 18 0.73 [0.25; 2.19] 0.5%
Langhorne, 2017 4 12 0.74 [0.31; 1.72] 0.7%
Nieto Estrada, 2017 3 12 0.74 [0.33; 1.63] 0.8%
Kreijkamp−Kaspers, 2017 4 5 0.74 [0.22; 2.54] 0.4%
Cluett, 2018 4 3 0.76 [0.36; 1.60] 0.9%
Hill, 2017 2 3 0.77 [0.08; 7.59] 0.1%
Beyar−Katz, 2017 5 3 0.79 [0.37; 1.70] 0.9%
Middleton, 2018 8 4 0.81 [0.13; 4.97] 0.2%
Vietto, 2018 7 5 0.82 [0.53; 1.27] 1.9%
Bauer, 2018 2 2 0.84 [0.30; 2.36] 0.5%
Ma, 2017 4 30 0.84 [0.33; 2.12] 0.6%
Goh, 2018 1 9 0.84 [0.03; 22.55] 0.1%
Legg, 2017 3 2 0.85 [0.46; 1.57] 1.2%
Zhu, 2018 1 3 0.85 [0.04; 19.17] 0.1%
Desborough, 2017 16 7 0.85 [0.51; 1.42] 1.5%
Ellis, 2017 7 9 0.86 [0.67; 1.09] 3.2%
Pinto−Sanchez, 2018 4 14 0.87 [0.46; 1.62] 1.2%
Chamberlain, 2017 18 12 0.87 [0.56; 1.35] 1.9%
Rodakowski, 2017 3 4 0.87 [0.61; 1.25] 2.3%
Wiffen, 2017 2 13 0.88 [0.47; 1.65] 1.1%
Wiysonge, 2017 1 4 0.88 [0.45; 1.72] 1.0%
Dvirnik, 2018 6 14 0.88 [0.44; 1.78] 1.0%
Macherey, 2017 1 2 0.88 [0.16; 4.89] 0.2%
Schandelmaier, 2017 3 9 0.89 [0.74; 1.06] 3.6%
Alfirevic Z, 2017 14 2 0.90 [0.32; 2.53] 0.5%
Lemiengre, 2018 5 7 0.92 [0.63; 1.35] 2.2%
McFadden, 2017 13 40 0.92 [0.72; 1.17] 3.2%
Stern, 2017 2 9 0.92 [0.22; 3.84] 0.3%
Gluud, 2017 9 3 0.93 [0.37; 2.30] 0.6%
Demicheli, 2018 11 14 0.93 [0.63; 1.38] 2.1%
Ohlsson, 2017 6 12 0.93 [0.54; 1.60] 1.4%
Thomas, 2018 10 7 0.94 [0.69; 1.28] 2.6%
Jefferson, 2018 4 4 0.94 [0.27; 3.22] 0.4%
Feinberg, 2017 62 55 0.96 [0.76; 1.21] 3.2%
Li, 2018 3 2 0.96 [0.06; 14.47] 0.1%
Cazeiro, 2017 3 6 0.96 [0.12; 7.93] 0.1%
Zhao, 2017 1 8 1.00 [0.63; 1.58] 1.8%
Alfirevic Z, 2017 3 9 1.03 [0.53; 2.01] 1.0%
Norman, 2017 5 8 1.04 [0.48; 2.22] 0.9%
Rice, 2017 15 29 1.04 [0.74; 1.46] 2.4%
Lee, 2018 1 3 1.04 [0.54; 2.01] 1.1%
Lenferink, 2017 1 13 1.05 [0.17; 6.59] 0.2%
Wong, 2017 1 3 1.06 [0.50; 2.23] 0.9%
Xu, 2018 3 9 1.06 [0.50; 2.23] 0.9%
Roberts, 2017 11 3 1.09 [0.52; 2.26] 0.9%
Hines, 2018 2 2 1.09 [0.00; 1071.86] 0.0%
Narula, 2018 8 3 1.09 [0.39; 3.03] 0.5%
Eke, 2017 2 27 1.11 [0.32; 3.81] 0.4%
Chauhan, 2017 1 3 1.11 [0.34; 3.67] 0.4%
Bighelli, 2018 20 17 1.12 [0.81; 1.57] 2.5%
Yao, 2017 2 3 1.13 [0.73; 1.74] 1.9%
Shepherd, 2017 1 3 1.14 [0.12; 10.91] 0.1%
Ruschen, 2018 2 2 1.17 [0.18; 7.73] 0.2%
Karam, 2017 7 5 1.19 [0.72; 1.98] 1.6%
Long, 2018 2 1 1.20 [0.03; 54.64] 0.0%
Kunutsor, 2017 2 21 1.21 [0.73; 2.01] 1.5%
Delara, 2018 2 6 1.21 [0.48; 3.04] 0.6%
Stacey, 2017 2 15 1.27 [0.40; 4.09] 0.4%
Toews, 2018 20 4 1.27 [0.34; 4.76] 0.3%
Demicheli, 2018 4 1 1.28 [0.64; 2.57] 1.0%
Walters, 2017 1 5 1.29 [0.23; 7.16] 0.2%
Laurant, 2018 3 4 1.31 [0.52; 3.35] 0.6%
Brown, 2017 1 9 1.36 [0.57; 3.27] 0.7%
Kaufman, 2018 1 5 1.38 [0.74; 2.55] 1.2%
Yin, 2017 2 8 1.39 [0.75; 2.58] 1.2%
Goldenburg, 2017 3 23 1.41 [0.48; 4.19] 0.5%
Offringa, 2017 5 3 1.45 [0.57; 3.65] 0.6%
Campbell, 2017 4 1 1.50 [0.29; 7.75] 0.2%
Liu, 2017 1 5 1.55 [0.67; 3.56] 0.7%
Vogel, 2017 3 2 1.57 [0.39; 6.42] 0.3%
Eamer, 2018 1 4 1.60 [0.77; 3.33] 0.9%
Teo, 2017 3 2 1.62 [0.11; 23.73] 0.1%
Gaskell, 2017 7 1 1.70 [0.52; 5.59] 0.4%
Hofmeyr, 2018 1 3 1.82 [0.97; 3.40] 1.2%
Jarde, 2017 1 12 1.96 [0.69; 5.51] 0.5%
Caissutti, 2017 2 13 2.02 [0.77; 5.33] 0.6%
Pattinson, 2017 4 1 2.02 [0.80; 5.11] 0.6%
Jiang, 2017 5 3 2.11 [1.04; 4.29] 1.0%
Wang, 2017 2 1 2.13 [0.53; 8.59] 0.3%
Sothornwit, 2017 1 2 2.15 [0.29; 15.99] 0.1%
Kahn, 2018 1 2 2.27 [0.78; 6.60] 0.5%
Roberge, 2017 10 8 2.34 [1.35; 4.07] 1.4%
Bryant−Smith, 2018 2 1 2.60 [0.57; 11.87] 0.2%
Woodley, 2017 1 5 2.95 [0.51; 17.16] 0.2%
Onwumeh, 2017 1 5 3.16 [0.12; 84.35] 0.1%
Middleton, 2017 7 2 3.64 [1.57; 8.45] 0.7%
Alabed S, 2018 4 2 4.05 [0.67; 24.48] 0.2%
Bromham, 2017 3 2 5.27 [1.82; 15.25] 0.5%
Gillies, 2017 10 1 10.71 [0.41; 282.23] 0.1%
Bjelakovic, 2017 1 2 14.63 [0.18; 1164.55] 0.0%
Random effects model 0.92 [0.85; 1.00] 100.0%
Heterogeneity: I^2 = 22%, τ^2 = 0.0350, p = 0.02
0.1
0.2
0.5
1
2
5
10
Before 2000 more beneficial
After 2000 more beneficial



Comparison	Q3 vs Q4	Q2 vs Q4	Q1 vs Q4
Ratio of odds ratios (95% CI)	0.98 (0.91 to 1.05)	0.89 (0.83 to 0.96)	0.85 (0.79 to 0.92)
τ^2	0.11	0.13	0.12


Adjustment on :
Number of contributive meta−analyses
ROR
CI (95%)
Random sequence generation
102
0.91
[0.84; 1.00]
Allocation concealment
96
0.92
[0.85; 1.00]
Blinding of participants and personnel
67
0.90
[0.84; 0.98]
Blinding of outcome assessment
79
0.91
[0.84; 1.00]
Incomplete outcome data
91
0.93
[0.86; 1.01]
Sample size
125
0.92
[0.86; 0.98]
0.9
1
1.1
Before 2000 more beneficial
Before 2000 less beneficial