

RESEARCH

Influence of trial sample size on treatment effect estimates: meta-epidemiological study

 OPEN ACCESS

Agnes Dechartres *assistant professor of epidemiology*^{1 2 3}, Ludovic Trinquart *senior statistician*^{1 4}, Isabelle Boutron *professor of epidemiology*^{1 2 3 4}, Philippe Ravaud *professor of epidemiology and director*^{1 2 3 4 5}

¹INSERM, U738, Paris, France; ²AP-HP (Assistance Publique des Hôpitaux de Paris), Hôpital Hôtel Dieu, Centre d'Épidémiologie Clinique, 75004 Paris, France; ³Univ Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris, France; ⁴French Cochrane Centre, Paris, France; ⁵Columbia University, Mailman School of Public Health, New York, NY, USA

Abstract

Objective To assess the influence of trial sample size on treatment effect estimates within meta-analyses.

Design Meta-epidemiological study.

Data sources 93 meta-analyses (735 randomised controlled trials) assessing therapeutic interventions with binary outcomes, published in the 10 leading journals of each medical subject category of the *Journal Citation Reports* or in the *Cochrane Database of Systematic Reviews*.

Data extraction Sample size, outcome data, and risk of bias extracted from each trial.

Data synthesis Trials within each meta-analysis were sorted by their sample size: using quarters within each meta-analysis (from quarter 1 with 25% of the smallest trials, to quarter 4 with 25% of the largest trials), and using size groups across meta-analyses (ranging from <50 to ≥1000 patients). Treatment effects were compared within each meta-analysis between quarters or between size groups by average ratios of odds ratios (where a ratio of odds ratios less than 1 indicates larger effects in smaller trials).

Results Treatment effect estimates were significantly larger in smaller trials, regardless of sample size. Compared with quarter 4 (which included the largest trials), treatment effects were, on average, 32% larger in trials in quarter 1 (which included the smallest trials; ratio of odds ratios 0.68, 95% confidence interval 0.57 to 0.82), 17% larger in trials in quarter 2 (0.83, 0.75 to 0.91), and 12% larger in trials in quarter 3 (0.88, 0.82 to 0.95). Similar results were obtained when comparing treatment effect estimates between different size groups. Compared with trials of 1000 patients or more, treatment effects were, on average, 48% larger in trials

with fewer than 50 patients (0.52, 0.41 to 0.66) and 10% larger in trials with 500-999 patients (0.90, 0.82 to 1.00).

Conclusions Treatment effect estimates differed within meta-analyses solely based on trial sample size, with stronger effect estimates seen in small to moderately sized trials than in the largest trials.

Introduction

Sample size varies greatly among trials, ranging from tens of patients to thousands of patients,¹ even within a meta-analysis investigating the same question. For example, a meta-analysis in cardiology² included trials with sizes ranging from 62 patients to 45 852 patients. Our knowledge about the influence of trial sample size on treatment effect estimates is based on the small study effect—the tendency for small trials to report greater treatment benefits than large trials in the same meta-analysis.³⁻⁵ A study based on a collection of meta-analyses in osteoarthritis showed that trials including fewer than 100 patients per arm yielded, on average, greater treatment effect estimates than did larger trials.⁶

The concept of a single threshold to distinguish small trials from large trials, whatever the medical area or intervention being tested, is not straightforward.⁷ For binary outcomes, the required sample size depends on the magnitude of treatment effect as well as the number of events and frequency of the medical condition. Therefore, a trial of 1000 patients can be considered large for certain medical conditions and small for others.

Correspondence to: A Dechartres agnes.dechartres@htd.aphp.fr

Extra material supplied by the author (see <http://www.bmj.com/content/346/bmj.f2304?tab=related#webextra>)

Web appendix 1: Selection of meta-analyses within each collection

Web appendix 2: Details concerning the statistical analyses performed

Web appendix 3: Characteristics of included meta-analyses and randomised controlled trials (RCTs) in the 2 collections of meta-analyses

Web appendix 4: Adjusted sensitivity analyses showing comparison of treatment effect estimates between quarters of sample size within meta-analyses and fixed thresholds

In this study, we assessed the influence of trial sample size on treatment effect estimates in a large collection of meta-analyses of various medical conditions and interventions.

Methods

Data sources

This study combined data from two independent collections of meta-analyses of randomised controlled trials assessing therapeutic interventions with binary outcomes. The first collection included 48 meta-analyses (421 trials) published in the 10 leading journals of each medical subject category of the *Journal Citation Reports* during two periods: between July 2008 and January 2009 and between January and June 2010 or in issue 4 of the *Cochrane Database of Systematic Reviews* in 2008. Further details about the search strategy and selection of meta-analyses have been published.⁸ Reports of all component trials from included meta-analyses were obtained.

The second collection included 45 meta-analyses (314 trials) published in the *Cochrane Database of Systematic Reviews* between January and July 2011. We included meta-analyses assessing a binary outcome for the primary or main outcome measure and involving four or more trials. If the meta-analysis reported combined results for more than one primary binary outcome, we selected the first reported outcome if it was described in enough trial reports. We excluded meta-analyses that overlapped with the first collection. Web appendix 1 includes details of the selection process for both collections.

Data extraction and risk of bias assessment

Using a standardised data extraction form, we extracted the following data for each randomised controlled trial: the date of publication, whether the trial was a single centre or multicentre trial (with at least two different centres), the number of patients with the outcome in each group, and the number of patients randomised in each group. Data for risk of bias were also collected by using the following domains of the risk of bias tool of the Cochrane Collaboration^{9 10}: methods for sequence generation and allocation concealment, blinding, and incomplete outcome data. Each domain was rated as having low, high, or unclear risk of bias, according to the recommendations of the Cochrane Collaboration.^{9 10} For each trial, the overall risk of bias was classified as low (that is, low risk of bias for all domains), high (that is, high risk of bias for one or more domains), or unclear (that is, unclear risk of bias for one or more domains in the absence of high risk of bias). We extracted data from the original reports of trials for the first collection of meta-analyses (in duplicate for a third of the meta-analyses), and from the Cochrane reviews for the second collection.

Data synthesis and analysis

Association between trial sample size and treatment effect

The trials within each meta-analysis were sorted by their sample size: using quarters within each meta-analysis (from quarter 1 including 25% of the smallest trials, to quarter 4 including 25% of the largest trials), and using size groups across meta-analyses (<50, 50-99, 100-199, 200-499, 500-999, and ≥ 1000 patients). Treatment effects (measured as odds ratios) were compared between quarters and size groups by multilevel logistic regression models with random effects.¹¹ These hierarchical models allowed for taking into account random intervention effects (between trial heterogeneity) within meta-analyses as well as random variation in the effect of trial sample size

between meta-analyses. The results were expressed as average ratios of odds ratios. This measure is the ratio of the odds ratio in smaller trials to the odds ratio in larger trials. A ratio of odds ratios less than 1 indicates larger estimates of the treatment effect in smaller trials. The heterogeneity across meta-analyses was quantified with τ^2 , the variance between meta-analyses. We performed tests for linear trend across quarters and size groups.

Sensitivity analyses

We reassessed the influence of trial sample size on treatment effect estimates by comparing treatment effects between trials by quarters (quarter 1 v quarters 2-4; quarters 1 and 2 v quarters 3 and 4; quarters 1-3 v quarter 4) and by fixed thresholds for trial sample size (50, 100, 200, 500, and 1000 patients; for example, for the 200 patient threshold, we compared treatment effects between trials with less than 200 patients and trials with 200 patients or more). We used the two stage approach to meta-epidemiological analyses as described by Sterne and colleagues¹² and further adjusted these analyses for the following trial characteristics: domains of risk of bias,¹³⁻¹⁷ overall risk of bias, centre status,^{8 18} and time since publication of the first trial within each meta-analysis.

Web appendix 2 details the statistical methods. We used SAS version 9.2 (SAS) for the multilevel models and Stata MP version 10.0 (Stata Corp) for the meta-epidemiological analyses.

Results

The study sample included 93 meta-analyses (735 randomised controlled trials; web appendix 3). A median of seven trials (range 3-30) were included per meta-analysis. Trial sample size varied greatly among the meta-analyses (median 34-2371 patients) and within the meta-analyses (for example, trial sample size ranged from 106 to 48 835 patients in one meta-analysis).

Association of trial sample size and treatment effect

Treatment effect estimates were significantly larger in smaller trials regardless of the sample size. Compared with trials in quarter 4 (which included the largest trials), treatment effects were, on average, 32% larger in trials in quarter 1 (which included the smallest trials; ratio of odds ratios 0.68, 95% confidence interval 0.57 to 0.82), 17% larger in trials in quarter 2 (0.83, 0.75 to 0.91), and 12% larger in trials in quarter 3 (0.88, 0.82 to 0.95). Heterogeneity across meta-analyses ranged from small to moderate in the three comparisons ($\tau^2=0.30, 0.07, \text{ and } 0.02$, respectively; fig 1⇓).

Compared with trials of 1000 patients or more, treatment effects were, on average, 48% larger in trials with fewer than 50 patients (ratio of odds ratios 0.52, 0.41 to 0.66), 34% larger in trials with 50-99 patients (0.66, 0.56 to 0.79), 30% larger in trials with 100-199 patients (0.70, 0.61 to 0.80), 19% larger in trials with 200-499 patients (0.81, 0.73 to 0.88), and 10% larger in trials with 500-999 patients (0.90, 0.82 to 1.00; fig 1). Heterogeneity across meta-analyses was moderate (τ^2 ranged from 0.11 to 0.26).

For both analyses, ratios of odds ratios showed a significant linear trend (both $P<0.001$).

Sensitivity analyses

In the two stage meta-epidemiological analyses, treatment effect estimates were, on average, 23% larger in quarter 1 trials (that is, the smallest trials) than in the other trials (ratio of odds ratios 0.77, 95% confidence interval 0.65 to 0.91), 19% larger in

quarter 1 and 2 trials than in quarter 3 and 4 trials (0.81, 0.74 to 0.88), and 15% larger in quarter 1-3 trials than in quarter 4 trials (that is, the largest trials; 0.85, 0.79 to 0.90). With comparisons of fixed thresholds of sample size, treatment effect estimates were also significantly larger in smaller trials, regardless of the threshold level. The heterogeneity across meta-analyses was low for all analyses (fig 2). Results were consistent after adjustment on the following trial characteristics: domains of risk of bias, overall risk of bias, centre status, and time of publication since the first trial (web appendix 4).

Discussion

In this meta-epidemiological study of 93 meta-analyses of 735 trials, we found significantly larger estimates of treatment effects in smaller trials, regardless of sample size. Treatment effect estimates differed within meta-analyses solely based on trial sample size, with, on average, stronger estimates in small to moderately sized trials than in the largest trials. The average difference was substantial and ranged from 12% to 32% when comparing estimates between quarters of sample size within meta-analyses.

Strengths and weaknesses of the study

Our results were based on a large meta-epidemiological study of 93 meta-analyses, representing various medical areas published in the leading journals of each medical speciality or in the *Cochrane Database of Systematic Reviews*. Cochrane reviews have generally been shown to be of higher methodological quality, are better reported, and have fewer conflicts of interest than do non-Cochrane reviews.¹⁹⁻²¹ To explore the influence of sample size on treatment effect, we used several complementary approaches, which all showed consistent results. However, because our results were based on meta-analyses of trials assessing binary outcomes, they cannot be extrapolated to trials assessing continuous outcomes because such trials usually differ in medical condition, risk of bias, sample size, and statistical analysis.

Possible explanations

Several mechanisms could help explain the association between trial sample size and treatment effects regardless of sample size. The first may be related to reporting bias. Smaller studies are more prone to publication bias,⁵ defined by the tendency for more likely publication of reports of studies with significant than non-significant results.²² A continuum of publication bias could exist to some extent: the larger the trial, the greater the probability that results are published, regardless of statistical significance. Smaller trials might also be more prone to outcome reporting bias.^{23 24}

Another possible mechanism is the difference in methodological quality by sample size.²⁵ Our results tended to be consistent after adjustment on domains of risk of bias, as well as overall risk of bias. Finally, it is possible that the larger the sample size, the greater the heterogeneity in selecting participants²⁶ or implementing interventions. Future research is needed to explore the effect of these different mechanisms.

Implications for researchers

Our results have an important effect on the interpretation of results of clinical trials and meta-analyses in general. The main issue of systematic reviews and meta-analyses is whether the combined treatment effect estimated by synthesising all included studies provides the best estimate of the true treatment effect,

or whether the studies overestimate or underestimate the treatment effect.²⁷ Because larger trials²⁸ are probably more pragmatic than smaller trials, with wider eligibility criteria²⁶ and greater variability in interventions, treatment effect estimates reported in large trials could be closer to the true treatment effect in real life. Thus, meta-analyses of all available evidence—whatever the trial sample size—might not reflect the true treatment effect.

Several authors have suggested that the results from large randomised controlled trials are inherently superior to those from smaller trials, even when the results of smaller trials are pooled in a meta-analysis,^{29 30} and that a substantially large trial should be conducted to definitively answer the question across a large sample of the population. Glasziou and colleagues³¹ also proposed relying on the results of the most precise trial if a meta-analysis was not available. Our results raise questions about whether the meta-analysis should be restricted to larger trials (or even to the “largest” trial). The downside of this approach would be imprecise estimates of the treatment effect. Rücker and colleagues recently proposed a method of limit meta-analysis, which allows for predicting treatment effects when the precision of each trial is increased to infinity.^{32 33} This approach is close to the regression based model described by Moreno and colleagues.³⁴

Conclusions

With a large meta-epidemiological study, we found smaller trials to have significantly larger estimates of treatment effects, regardless of sample size. Effect estimates differed within meta-analyses solely based on trial sample size, with, on average, stronger effect estimates in small to moderately sized trials than in the largest trials. These stronger effects might not reflect the true treatment effect; therefore, robustness of the conclusions of a meta-analysis, including assessment of the influence of trial sample size on treatment effect estimates using sensitivity analyses (for example, subgroup analyses comparing quarters or limit meta-analysis), should be assessed. Reviewers and readers can easily check whether the result for the overall meta-analysis agrees with the results for the largest trials (that is, those in quarter 4 of sample size). Interpretation of the pooled result should be cautioned when this is not the case. More generally, our results raise questions about how meta-analyses are currently performed, especially whether all available evidence should be included in meta-analyses because it could lead to more beneficial results.

We thank Mickaël Randrianandrasana for data extraction from the Cochrane reports for the second collection of meta-analyses.

Contributors: AD is the study guarantor, and was involved in the study conception, selection of trials, data extraction, data analysis, interpretation of results, and drafting of the manuscript; LT was involved in the study conception, data analysis, interpretation of results, drafting of the manuscript; IB and PR were involved in the study conception, interpretation of results, and drafting of the manuscript. All authors, external and internal, had full access to all the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This study was funded by an academic grant from the Programme Hospitalier de recherche Clinique Régional (AOR10017). Our team is supported by an academic grant (DEQ20101221475) for the programme “Equipe espoir de la Recherche,” from the Fondation pour la Recherche Médicale. The researchers declare that they are independent from the funders. The funders did not have any role in the study.

What is already known on this topic

- Sample size varies greatly among trials, ranging from tens of patients to thousands of patients, even within one meta-analysis answering the same question
- Small study effect has been previously defined as the tendency for small trials in a meta-analysis to show larger treatment benefits
- The concept of using a single threshold to distinguish small and large trials, whatever the medical area or intervention being tested, is not straightforward

What this study adds

- Smaller trials had significantly larger estimates of treatment effect, regardless of sample size.
- Effect estimates differed within meta-analyses solely based on trial sample size, with, on average, stronger estimates seen in small to moderately sized trials than in the largest trials
- The robustness of the conclusions of a meta-analysis should be assessed with careful interpretation of results if the overall result is not consistent with those of the largest trials

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: support from the Programme Hospitalier de recherche Clinique Régional and the Fondation pour la Recherche Médicale for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

- Hopewell S, Dutton S, Yu LM, Chan AW, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ* 2010;340:c723.
- Brandler E, Paladino L, Sinert R. Does the early administration of beta-blockers improve the in-hospital mortality rate of patients admitted with acute coronary syndrome? *Acad Emerg Med* 2010;17:1-10.
- Sterne JA, Egger M, Smith GD. Systematic reviews in health care: investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;323:101-5.
- 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;53:1119-29.
- Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002.
- Nuesch E, Trelle S, Reichenbach S, Rutjes AW, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ* 2010;341:c3515.
- Ioannidis JP, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA* 1998;279:1089-93.
- Dechartres A, Boutron I, Trinquart L, Charles P, Ravaud P. Single-center trials show larger treatment effects than multicenter trials: evidence from a meta-epidemiologic study. *Ann Intern Med* 2011;155:39-51.
- Higgins JP, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
- Higgins JPT, Green S, eds. Cochrane handbook for systematic reviews of interventions, version 5.0.0. Cochrane Collaboration, 2008.
- 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;26:2745-58.
- 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;21:1513-24.
- 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;352:609-13.
- 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.
- 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;36:847-57.
- 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;273:408-12.
- 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;336:601-5.
- Bafeta A, Dechartres A, Trinquart L, Yavchitz A, Boutron I, Ravaud P. Impact of single centre status on estimates of intervention effects in trials with continuous outcomes: meta-epidemiological study. *BMJ* 2012;344:e813.
- Moja LP, Telaro E, D'Amico R, Moschetti I, Coe L, Liberati A. Assessment of methodological quality of primary studies by systematic reviews: results of the metaquality cross sectional study. *BMJ* 2005;330:1053.
- Moseley AM, Elkins MR, Herbert RD, Maher CG, Sherrington C. Cochrane reviews used more rigorous methods than non-Cochrane reviews: survey of systematic reviews in physiotherapy. *J Clin Epidemiol* 2009;62:1021-30.
- Shea B, Moher D, Graham I, Pham B, Tugwell P. A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. *Eval Health Prof* 2002;25:116-29.
- Egger M, Smith GD. Bias in location and selection of studies. *BMJ* 1998;316:61-6.
- Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004;291:2457-65.
- Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010;340:c365.
- Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med* 2001;135:982-9.
- Yusuf S, Held P, Teo KK, Toretzky ER. Selection of patients for randomized controlled trials: implications of wide or narrow eligibility criteria. *Stat Med* 1990;9:73-83; discussion 83-6.
- Goodman S, Dickersin K. Metabias: a challenge for comparative effectiveness research. *Ann Intern Med* 2011;155:61-2.
- Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? *Stat Med* 1984;3:409-22.
- Flather MD, Farkouh ME, Pogue JM, Yusuf S. Strengths and limitations of meta-analysis: larger studies may be more reliable. *Control Clin Trials* 1997;18:568-79; discussion 661-6.
- LeLorier J, Gregoire G, Benhaddad A, Lapierre J, Derderian F. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997;337:536-42.
- Glasziou PP, Shepperd S, Brassey J. Can we rely on the best trial? A comparison of individual trials and systematic reviews. *BMC Med Res Methodol* 2010;10:23.
- Rücker G, Carpenter JR, Schwarzer G. Detecting and adjusting for small-study effects in meta-analysis. *Biom J* 2011;53:351-68.
- Rücker G, Schwarzer G, Carpenter JR, Binder H, Schumacher M. Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistics* 2011;12:122-42.
- Moreno SG, Sutton AJ, Thompson JR, Ades AE, Abrams KR, Cooper NJ. A generalized weighting regression-derived meta-analysis estimator robust to small-study effects and heterogeneity. *Stat Med* 2012;31:1407-17.

Accepted: 2 April 2013

Cite this as: *BMJ* 2013;346:f2304

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Figures

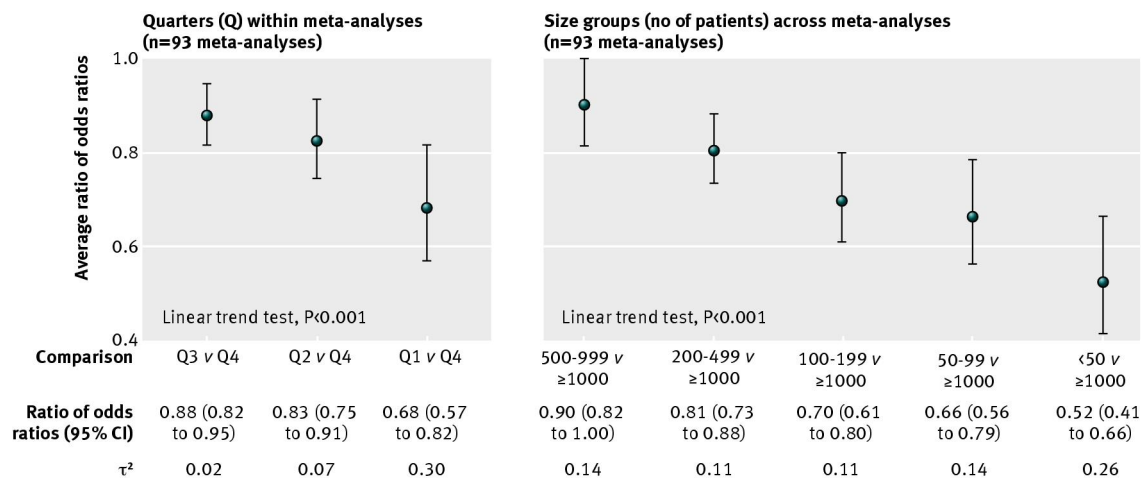


Fig 1 Comparison of treatment effect estimates between trial sample sizes, grouped by quarters (from quarter 1 with the smallest trials, to quarter 4 with the largest trials) and by size groups (from <50 patients to ≥1000 patients). Both analyses included all 93 individual meta-analyses (735 randomised controlled trials)

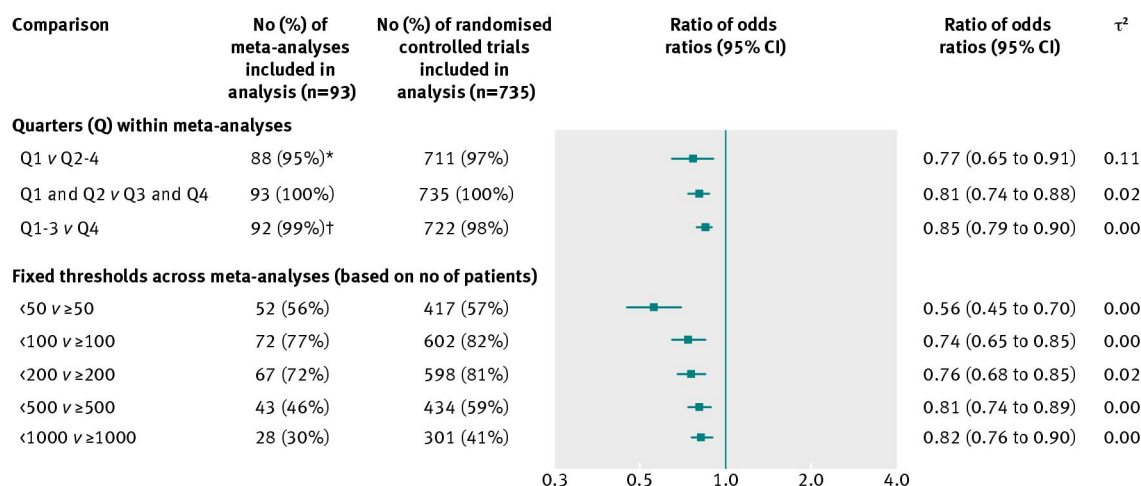


Fig 2 Comparison of treatment effect estimates between trial sample sizes grouped by quarters (from quarter 1 with the smallest trials, to quarter 4 with the largest trials) and by fixed thresholds (according to numbers of patients). All 93 meta-analyses did not contribute to the analysis, depending on the threshold used (at least one trial with a sample size less and more than the threshold within each meta-analysis was required to perform these analyses). *Five meta-analyses did not contribute to the analysis (three included fewer than four trials; for two trials, the meta-regression model did not converge, therefore, the ratio of odds ratios could not be obtained). †One meta-analysis did not contribute to the analysis (the meta-regression model did not converge, therefore, the ratio of odds ratios could not be obtained)